

1561

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN GÉNIE INDUSTRIEL

PAR
PHILIPPE POIRIER

PRODUCTION JAT À L'AIDE DE LA MÉTHODE CONWIP:
ÉTUDE COMPARATIVE PAR SIMULATION

Août 1997

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

RÉSUMÉ

Les études comparatives par simulation effectuées dans le domaine du Juste-À-Temps (JAT) utilisent essentiellement des systèmes de production composés de lignes d'assemblage ou des systèmes simplifiés de type atelier spécialisé. Le but de cette recherche est d'étudier le comportement, dans un environnement JAT, d'un système réel de production relevant du domaine du meuble et bois ouvré. Le système à l'étude est considéré dans son ensemble. Il est composé d'une partie atelier spécialisé qui alimente des lignes d'assemblage.

La méthode CONWIP est utilisée sous une gestion en flux tirés afin d'adapter ce système au JAT. Cette méthode est une extension de la méthode KANBAN qui repose essentiellement sur deux points: (a) l'emploi de flux de production hybrides (poussés et tirés); et (b) la détermination des tailles de lots de production est effectuée en fonction des goulots d'étranglement du système.

Après avoir établi par simulation une base de comparaison des flux poussés versus tirés à l'aide des Méthodes TAGUCHI, les simulations de validation sont utilisées pour l'analyse

comparative. Il en résulte que si l'on utilise le JAT pour ce système, il ne faut pas nécessairement espérer une diminution du temps de passage des commandes usine. Le gain est cependant significatif lorsque l'on observe le niveau d'en-cours nécessaire au bon fonctionnement de l'usine qui est moindre.

REMERCIEMENTS

J'aimerais remercier sincèrement mes directeurs de recherche, soit Messieurs René Rochette et Georges Abdounour pour leur soutien et leur disponibilité. J'aimerais également remercier le Conseil National en Science Naturelle et en Génie du Canada ainsi que le Laboratoire de Recherche en Productique de l'Université du Québec à Trois-Rivières pour leur aide financière soutenue. Sans elle, cette recherche n'aurait jamais eu lieu.

TABLE DES MATIÈRES

	Page
RÉSUMÉ.....	ii
REMERCIEMENTS.....	iv
LISTE DES TABLEAUX.....	vi
LISTE DES FIGURES.....	viii
CHAPITRES	
I. INTRODUCTION.....	1
La définition du problème.....	1
Les objectifs de cette recherche.....	7
L'utilité de tirer les flux.....	8
Adapter le JAT à une variété d'environnements.....	12
II. LA FORMULATION DU MODÈLE.....	23
Le système de production à l'étude.....	23
Les hypothèses émises.....	23
Formulation mathématique.....	37
Le design expérimental.....	50
III. LE SIMULATEUR ET LA VALIDATION DU MODÈLE.....	60
Le simulateur.....	60
La validation.....	64
IV. RÉSULTATS: ANALYSE ET DISCUSSION.....	69
Résultats en flux poussés.....	70
Résultats en flux tirés.....	74
La validation des modèles et l'analyse comparative.....	78
V. CONCLUSIONS ET RECOMMANDATIONS.....	87
RÉFÉRENCES.....	90
ANNEXE A. Modélisation graphique du réseau en SLAM II.....	94
ANNEXE B. Le facteur de sécurité.....	111

LISTE DES TABLEAUX

	Page
1. Liste et description des règles d'ordonnement utilisées.....	49
2. Les paramètres expérimentaux utilisés pour toutes les simulations.....	53
3. Méthodes TAGUCHI: table orthogonale $L_{16}(2^{15})$	54
4. Facteurs expérimentaux en flux poussés.....	55
5. Facteurs expérimentaux en flux tirés.....	56
6. Assignation systématique des niveaux pour les facteurs en flux poussés.....	57
7. Assignation systématique des niveaux pour les facteurs en flux tirés.....	58
8. Assignation des règles d'ordonnement locales.....	59
9. Résultats de simulation en flux poussés: temps de passage moyen des commandes usine.....	70
10. Résultats de simulation en flux poussés: temps de passage moyen des commandes usine (moyenne et écart- type).....	71
11. Résultats de simulation en flux poussés: temps de passage moyen des commandes usine (analyse de la variance).....	72
12. Temps de passage moyen des commandes usine en flux poussés: régression (modèle ajusté).....	74
13. Résultats de simulation en flux tirés: temps de passage moyen des commandes usine.....	75
14. Résultats de simulation en flux tirés: temps de passage moyen des commandes usine (moyenne et écart- type).....	75

	Page
15. Résultats de simulation en flux tirés: temps de passage moyen des commandes usine (analyse de la variance).....	76
16. Temps de passage moyen des commandes usine en flux tirés: régression (modèle ajusté).....	78
17. Temps de passage moyen des commandes usine en flux poussés: test de Duncan.....	80
18. Temps de passage moyen des commandes usine en flux tirés: test de Duncan.....	81
19. Affectation des niveaux utilisés lors des essais de validation en flux poussés et tirés.....	82
20. Validation en flux poussés: compilation des résultats	83
21. Validation en flux tirés: compilation des résultats..	83
22. Résultats de l'étude comparative.....	86

LISTE DES FIGURES

	Page
1. Schématique du système productif modélisé.....	24
2. Les tables de données reliées au système de production à l'étude.....	51
3. Vue globale du progiciel de simulation.....	61
4. Transition entre les régimes transitoire et permanent observée pour une simulation donnée.....	63
5. Temps de passage moyen des commandes usine en flux poussés: simulations de validation (intervalle de confiance à 95%).....	84
6. Temps de passage moyen des commandes usine en flux tirés: simulations de validation (intervalle de confiance à 95%).....	84
7. Les blocs ressources.....	95
8. L'aiguilleur du réseau.....	96
9. Les branches reliées à l'avancement du temps.....	98
10. La gestion intercentre de travail.....	100
11. La gestion du niveau physique des stocks.....	102
12. Le regroupement des composants avant l'assemblage....	103
13. Les ressources d'assemblage.....	104
14. Les postes de travail de l'atelier spécialisé.....	106
15. L'éclatement des commandes clients et leur reconsti- tution.....	110
16. Le facteur de sécurité utilisé pour la gestion des stocks: exemple.....	112

CHAPITRE I

INTRODUCTION

La définition du problème

L'approche Euro-américaine de planification de la production la plus répandue est sans aucun doute la planification des besoins en composants (PBC), connue sous le vocable de MRP ("Material Requirements Planning"). La première version de cette approche remonte aux années 1960. Elle visait à établir une programmation de la production, tout en considérant le système de production à capacité infinie. L'optimisation de l'utilisation des facteurs productifs était donc effectuée dans un contexte où toutes les propositions faites devaient être supposément réalisables.

Il a fallu attendre la fin des années 1970 pour voir apparaître une version qui cherchait à ajuster la charge souhaitée à la charge disponible pour chaque poste de charge. Connue sous le nom de management des ressources de la production, ou MRP2 ("Manufacturing Resources Planning"), cette version n'a pas cessé d'évoluer. Aujourd'hui, certaines entreprises sont dotées de logiciels personnalisés qui traitent

toutes les informations relatives aux ressources impliquées dans leurs systèmes.

Malgré son coût prohibitif et sa complexité, cette gestion de production centralisée et assistée par ordinateur semblait être la voie la plus prometteuse pour faire face à une concurrence accrue, que la mondialisation des marchés n'a fait que renforcer.

Cependant, un article publié dans les années 1970 (Sugimori et al. 1977) portant sur l'implantation du Juste-À-Temps (JAT) dans une entreprise ébranlera cette conception futuriste. Depuis, l'engouement pour comprendre et étudier la philosophie du JAT n'a d'égal que le nombre de publications sur ce sujet. Cette ferveur est attribuable d'une part au succès inattendu des systèmes de production japonais, et d'autre part à la dualité de cette nouvelle approche de gestion de production (en relation avec l'approche MRP).

Le contraste le plus frappant est sans aucun doute que le JAT utilise un mode de gestion décentralisé des flux d'information et des flux de production. Son médium est le Kanban, qui n'est que la composante manuelle de la gestion des flux. L'utilisation de Kanbans simplifie la gestion opérationnelle. Ils permettent de réduire les en-cours,

régulariser les flux de production et ils facilitent le contrôle de la production. De plus, le jalonnement et l'ordonnancement des opérations s'en trouvent simplifiés.

Traditionnellement, les flux de production d'un système de type MRP sont "poussés". À partir d'une programmation prévisionnelle anticipant la demande en composants, on pousse ces derniers de poste en poste, de l'amont vers l'aval. Ainsi, on contrôle le taux de production du système, tout en mesurant les volumes d'en-cours obtenus, selon Spearman et Zazanis (1992). En conservant des en-cours volumineux, on peut minimiser les impacts des aléas sur la production.

Inversement, la production JAT suit une logique de production à flux tirés. La production d'un composant est amorcée par la demande des postes de charge demandeurs de la référence. Le JAT utilise, tout comme la MRP, une sorte de programmation prévisionnelle. Par contre, le système est contrôlé par les volumes d'en-cours. C'est le taux de production qui sera mesuré afin d'obtenir une certaine rétroaction.

La production JAT est plus qu'un simple mode de gestion des flux. C'est une philosophie dont le fondement est l'élimination de toutes les sources de gaspillage dans la

production en fournissant le bon composant en quantité nécessaire au bon moment. Pour y arriver, Monden (1991) affirme dans son ouvrage relié au système de production Toyota qu'il est préférable pour cette entreprise de gérer sa production en flux tirés. D'une façon plus générale, Wantuck (1989) définit le JAT comme une stratégie de production englobant un nouvel ensemble de valeurs qui favorise l'amélioration continue de la qualité et de la productivité.

Même s'il est généralement admis que dans un environnement JAT, philosophie et flux tirés vont de pair, il n'existe qu'un faible consensus sur l'importance relative de certaines variables nécessaires à l'implantation du JAT. Golhar et Stamm (1991) ont passé en revue la littérature sur ce sujet, et ils ont identifié quatre principes qui sont gages de succès pour l'implantation d'une telle philosophie: (a) élimination du gaspillage; (b) participation des employés aux décisions; (c) participation des fournisseurs; et (d) maîtrise de la qualité.

Même si l'application des trois derniers principes offre des avantages non négligeables, ils ne feront pas l'objet de cette étude. Nous nous concentrerons sur l'essence même du JAT.

Dans son ouvrage portant sur la gestion de production, Giard (1988) résume bien la hantise du gaspillage des Japonais.

Elle s'explique par leur exigüité géographique ainsi que par l'absence quasi-totale de ressources minières ou énergétiques. Ce qui ne contribue pas directement à ajouter de la valeur à un produit est considéré comme du gaspillage. La philosophie japonaise nous conduit donc à considérer les stocks comme du gaspillage. Car ceux-ci ne génèrent pas de valeur ajoutée. Ainsi, tout ce qui pourra contribuer à diminuer les en-cours (ou plus précisément les stocks-tampons pour une gestion en flux tirés) sera considéré comme une variable importante.

Mais puisque le stock sert de tampon entre l'unité productrice et l'unité demandeuse, une diminution de celui-ci conduit à une interdépendance des problèmes de gestion. Voilà pourquoi l'implantation du JAT nécessite la mise en application de plusieurs principes, comme l'amélioration continue et la maintenance préventive. Par contre, les avantages qui en résultent sont considérables.

L'élimination de la sécurité procurée par les stocks nous amène à identifier deux familles de variables critiques qui affectent la performance d'un système de production JAT: (a) les variables reliées à la fiabilité d'un système; et (b) les variables de flexibilité.

La première famille de variables inclut la variabilité des temps opératoires, les pannes, la synchronisation des flux

de production, etc... Les variables de la seconde famille sont plus subtiles. Un système ayant un minimum de stocks doit nécessairement être plus réactif pour faire face aux aléas et aux fluctuations de la demande. Il faut donc diminuer le temps de passage nécessaire à la fabrication d'un produit en diminuant les temps de mise en course et la taille des lots de production, augmenter la flexibilité de la main-d'oeuvre, recourir à des routages alternatifs, anticiper judicieusement la demande, etc...

Comme il a déjà été mentionné, l'importance relative de ces variables est difficile à établir. La quasi-globalité des études portant sur la performance ou la sensibilité d'un système JAT ont été effectuées en utilisant une modélisation correspondant à une ligne de production. Deux raisons peuvent expliquer ce constat. La première raison est qu'une ligne de production est facilement modélisable et que l'on peut généraliser aisément les résultats. La seconde est que plusieurs personnes croient, à tort, que le JAT est réservé à la production de masse dont la demande est hautement répétitive. Pourtant, les principes du JAT peuvent être bénéfiques pour une multitude d'environnements de production.

Les objectifs de cette recherche

Le but de cette recherche est d'étudier le comportement d'un système réel de production de type atelier spécialisé, qui produit en lots intermittents, dans un environnement de JAT. Vue la complexité de ce type de système, l'étude sera donc effectuée par simulation.

La justification de ce genre d'étude est évidente. Golhar et Stamm (1991) soulignent la pauvreté de la recherche dans ce domaine. Ils le qualifient même de secteur critique. Si tous s'entendent pour dire qu'il faut préalablement modifier le système en aménagement cellulaire, qu'il faut également utiliser les techniques du groupement analogique, il y a très peu de chercheurs qui ont exploré cette voie. La contribution de ce travail au génie industriel sera appréciable si l'on rencontre les objectifs suivants: (a) adapter le JAT à un système réel de production de type atelier spécialisé; (b) établir une base de comparaison entre un système conventionnel géré en flux poussés et un système géré en flux tirés ou hybrides; et (c) comparer ceux-ci.

Le premier objectif ne présente pas de difficulté particulière. Plusieurs scénarios ont été proposés dans différents articles, mais très peu ont été explorés. Généralement, les études se limitaient à une cellule de

fabrication. Ainsi, en travaillant sur un système industriel plus complexe, nous pourrions en tirer des enseignements intéressants et novateurs.

Le troisième objectif ne pose également pas de problème. Il y a eu suffisamment de publications permettant d'identifier facilement certains facteurs expérimentaux qui sont intéressants à étudier dans ce domaine.

Par contre, le second objectif est beaucoup plus délicat à traiter. Afin de comparer la performance d'un système fonctionnant dans un environnement JAT, il faut nécessairement avoir une base de comparaison. Même si la modélisation du système initial (le système actuel) est importante pour la validation, on ne peut se contenter de comparer celui-ci avec le nouveau système obtenu. Il faudra donc modéliser également un autre système similaire à celui-ci, mais fonctionnant en flux poussés.

L'utilité de tirer les flux

Nous savons que le JAT est une philosophie qui repose sur des principes que l'on peut qualifier intuitivement de valables. Cependant, on peut se questionner sur l'utilité de gérer la production en flux tirés.

Une question de fiabilité

Plusieurs chercheurs ont comparé la performance de systèmes fonctionnant en flux tirés et poussés. Par exemple, Sarker et Fitzsimmons (1989) ont étudié un système constitué d'une série de neuf postes de travail identiques, dont les temps opératoires étaient normalement distribués (la normale étant tronquée). Leurs travaux mettent en évidence la supériorité des flux poussés. Et ce, pour un même volume d'encours dans les deux systèmes. Ils concluent qu'un système en flux tirés est plus affecté par une variation des temps opératoires, et que le taux de production diminue non-linéairement et de manière drastique lorsque la variabilité augmente.

Par contre, Baker et al. (1990) doutent de ces résultats. Ils reprennent l'étude et ils obtiennent des résultats contradictoires. La supériorité apparente des flux poussés serait attribuable à l'utilisation de volumes d'encours plus élevés, comparativement à l'utilisation de faibles stocks-tampons lors des simulations en flux tirés. Ainsi, la comparaison serait biaisée. De plus, ils tracent des courbes du taux de production en fonction du taux de variation des temps opératoires pour plusieurs distributions: la normale tronquée, la lognormale, l'uniforme et l'Erlang. Non seulement ces courbes se superposent, mais elles diffèrent également de

celles de Sarker et Fitzsimmons. Donc, ceux-ci auraient faussé leurs résultats, selon Baker et al.

En fait, il est normal de penser qu'en réduisant les stocks, un système de production sera plus affecté par les aléas. Une vaste étude de comparaison a été effectuée par Krajewski et al. (1987). Ils démontrent que la performance d'un système "Kanban" est sensible à son environnement. Ainsi, un système ayant un haut taux de rejets, une main-d'oeuvre peu flexible, un taux de pannes élevé ainsi qu'un répartition inéquitable de la charge de travail sur une ligne de production serait moins performant qu'un système de type MRP, pour un même niveau d'en-cours.

Ils déclarent qu'un système Kanban n'est pas crucial en soi pour améliorer la performance, ce n'est qu'une composante d'une philosophie globale. L'important, c'est de travailler sur les facteurs qui rendront les flux de production plus uniformes ainsi que le système plus flexible, afin qu'il s'adapte mieux à la capacité requise. Dans la même veine, réduire les temps de mise en course et diminuer la taille des lots sont des objectifs efficaces pour diminuer les coûts reliés aux stocks et améliorer le service à la clientèle.

De ce cheminement, on retire qu'un système géré en flux tirés force les différents intervenants à améliorer leurs facteurs productifs. Nécessairement, ceci est vrai si les stocks-tampons sont réduits.

Une question de contrôle

Comme il a déjà été mentionné, pour gérer un système en flux poussés, on contrôle le taux de production en le considérant comme une variable indépendante et on mesure les en-cours. Tandis qu'en flux tirés, le contrôle est effectué sur les en-cours et on mesure le taux de production qui devient alors une variable dépendante. À priori, on peut difficilement identifier quel mode de gestion offre les meilleures opportunités.

En tenant compte de considérations pratiques, on peut penser premièrement qu'un contrôle sur les en-cours est plus aisé, puisque ceux-ci sont directement observables. Deuxièmement, un contrôle par le taux de production implique une estimation de la capacité. Si cette estimation est surévaluée, on obtiendra un surplus d'en-cours non désiré. De plus, cette estimation engendre un risque certain, puisqu'il faut compter avec les aléas.

Adapter le JAT à une variété d'environnements

Les modélisations hypothétiques représentant une ou des lignes de production sont de loin celles qui ont été le plus utilisées dans le domaine du JAT. À ce sujet, Chu et Shih (1992) se questionnent sur la transférabilité des résultats. Ils font remarquer notamment que la plupart des études par simulation supposent que le système de production est parfait. Même si certaines hypothèses qui en découlent reflètent bien les caractéristiques d'une production JAT, est-ce représentatif d'un système réel? On peut penser qu'une variable affectant négativement la performance d'un système hypothétique le fera également pour un système réel. Mais, que sait-on sur l'amplitude de l'effet?

À la lumière des résultats obtenus par diverses simulations, plusieurs personnes ont rapidement conclu que le JAT est applicable seulement pour un certain type d'environnement qui possède les caractéristiques suivantes: (a) production de masse; (b) temps de mise en course peu significatif; (c) demande hautement répétitive; et (d) bonne maîtrise de la qualité.

Ceci est vrai pour un système Kanban à une ou deux cartes qui est appliqué dans sa forme la plus pure. Mais pour un

système de production de type atelier spécialisé, en est-il de même?

Simulation de systèmes industriels divers

Très peu d'études par simulation ont été réalisées sur des systèmes que l'on peut qualifier de non-conventionnels. Ramudhin et Rochette (1991) ont modélisé un système réel comportant cinq postes de charge pouvant contenir chacun plus d'une machine en parallèle. Même si les temps de mise en course ainsi que les temps opératoires étaient déterministes, l'arrivée des commandes atelier était générée aléatoirement, selon une distribution représentant cinq familles de produits, chaque famille ayant son propre routage. Puisqu'il n'y avait pas de regroupement de commandes, ni de lissage de la production, la variation de la taille des lots de production était élevée. De plus, les temps entre deux pannes ainsi que les temps de réparation suivaient des distributions exponentielles variées.

La modélisation en flux poussés utilisée impliquait notamment des volumes d'en-cours illimités, et l'ordonnancement des opérations se résumait à allouer une règle de priorité locale à chaque poste de charge.

La modélisation en flux tirés différait quelque peu de la méthode Kanban, telle qu'on la connaît. Chaque machine avait un stock-tampon ainsi qu'une liste de requêtes provenant des autres machines. Lorsqu'une tâche était complétée à une machine, l'en-cours correspondant était acheminé vers la machine qui en avait fait la demande, si la capacité du stock-tampon de la machine demanderesse n'était pas excédée. Car si tel était le cas, la machine expéditrice était bloquée jusqu'à ce qu'un volume suffisant se libère. Après le transfert, si son volume de stock-tampon était inférieur à un niveau critique, on regardait la liste de requête pour débiter une nouvelle tâche, tout en envoyant également une requête à une des machines en amont (les flux étaient unidirectionnels). Si le niveau critique n'était pas dépassé, on se contentait de regarder la liste des requêtes pour débiter une nouvelle tâche, s'il y en avait de disponible.

De cette manière, si la capacité des stocks-tampons est égale à un, que le niveau critique est égal à zéro et que le nombre de requêtes à envoyer est unitaire, on obtient un système en flux tirés dans sa forme la plus pure. À l'inverse, si on augmente ces trois paramètres, le système tend vers une gestion en flux poussés.

Les résultats de cette étude de comparaison montrent que pour un même temps de passage dans le système, une gestion en flux tirés offre les avantages suivants: (a) il y a moins de commandes sur le plancher de l'usine; (b) les volumes d'en-cours sont significativement réduits; et (c) la variation du temps de passage sur le plancher de l'usine est moindre.

Ainsi, même si l'environnement d'un système de production est instable, un réseau de files d'attente fermé peut être supérieur à un réseau ouvert. Cette supériorité peut être attribuable en bonne partie au fait que le nombre d'ordres de fabrication pour un système en flux tirés est négativement corrélé par rapport au nombre de postes de travail, tandis qu'il ne l'est pas nécessairement pour un système en flux poussés. Cette notion de dépendance négative est discutée sommairement par Spearman *et al.* (1990).

Une autre étude de comparaison non-conventionnelle a été réalisée par Durmosoglu (1991). À partir d'un système industriel, il a appliqué les techniques du groupement analogique afin d'obtenir un environnement cellulaire. En procédant ainsi, il voulait profiter des avantages reliés à ces techniques: (a) réduction des temps de mise en course; (b) réduction du temps de passage sur le plancher de l'usine; (c)

diminution des volumes d'en-cours; et (d) simplification des routages.

Cependant, l'étude de comparaison s'est limitée à une seule cellule qui alimente une ligne d'assemblage. Là encore, les flux étaient unidirectionnels. De plus, les temps de mise en course ainsi que les temps opératoires étaient déterministes. Par contre, les temps d'assemblage suivaient une distribution triangulaire. Les règles de priorité utilisées pour la cellule correspondaient soit au premier-arrivé, premier-servi, soit à un chevauchement de certains routages afin de minimiser la congestion dans la cellule.

La modélisation en flux tirés différait également de la méthode Kanban conventionnelle. Au lieu de tirer les flux sur la dernière machine de la cellule, on envoyait l'étiquette de production à l'entrée de la cellule. Et les flux étaient poussés à l'intérieur de celle-ci.

Les résultats obtenus montrent que si les temps de mise en course ne peuvent être réduits, on peut obtenir des performances similaires pour les systèmes tiré et poussé. Cependant, le contrôle du système en flux tirés est simple et automatique. Tandis que pour le système en flux poussés, le mécanisme de contrôle est beaucoup plus complexe.

De plus, l'auteur rapporte qu'une diminution significative des temps de mise en course augmente considérablement la performance du système. Néanmoins, si cette diminution dépasse un certain seuil critique, son effet sera moins significatif par rapport à une bonne programmation des opérations. De plus, si cette diminution est accompagnée d'une baisse draconienne de la taille des lots de production et de transfert, il ne faudra pas espérer une augmentation de la performance (si le taux de production et le volume d'en-cours sont les critères de performance). Donc, si les temps de mise en course ne sont pas négligeables, il est préférable de travailler avec des lots économiques. Et ce, peu importe que le système soit tiré ou poussé.

La méthode CONWIP

Comme on a pu s'en rendre compte, il est possible d'obtenir de bons résultats en production JAT avec des systèmes ayant un environnement non-conventionnel par rapport à l'environnement idéal décrit antérieurement. Pour ce faire, il suffit d'adapter la production JAT aux particularités de l'environnement.

Dans un article portant sur de nouveaux concepts reliés à la planification et au contrôle de la production, Zäpfel et Missbauer (1993) présentent un système de contrôle de

la production similaire à la méthode Kanban, mais qui peu être applicable à une multitude d'environnements. C'est la méthode CONWIP ("CONstant Work-In-Process"). L'idée de base de cette méthode est d'établir un volume d'en-cours normalisé pour chaque routage, afin de conserver le même temps opératoire au goulot d'étranglement, et ce pour tous les composants.

Cette méthode a été présentée initialement par: (a) Spearman, Woodruff et Hopp (1990); et (b) Spearman et Zizanis (1992). CONWIP est une extension de la méthode Kanban. Considérons simplement une ligne de production à flux unidirectionnels, avec un stock-tampon à chaque poste de travail et dont les conteneurs sont normalisés en quantité. De plus, assumons que ce système est contrôlé en fonction des stocks (en flux tirés). Au lieu d'acheminer un Kanban au dernier poste de charge de la ligne, on envoie un signal de production au début de cette dernière. Les flux sont poussés à l'intérieur de celle-ci.

Le signal envoyé est générique. C'est-à-dire que lorsqu'il est envoyé, aucune demande spécifique en composant(s) n'est rattachée à ce signal. Si on utilise un système de cartes, on attachera une carte générique au début de la ligne à un conteneur. À la fin de la ligne, cette même carte sera réacheminée au début, en attente sur un babillard, pour être

éventuellement attachée à un autre conteneur de pièces. Lorsque le premier poste de charge a besoin d'être réalimenté, on assignera un numéro de composant(s) à une carte générique, en consultant la liste des commandes atelier spécifique à cette ligne de production.

Afin de déterminer la taille des lots de production de chaque composant, on identifiera le (ou les) goulot d'étranglement du système. La taille des lots sera donc fonction de la nomenclature du produit à fabriquer ainsi que du temps opératoire que l'on désire maintenir constant au goulot. Habituellement, il y a très peu de postes de charge qui constituent des goulots d'étranglement, et ceux-ci sont faciles à identifier. Conserver un temps opératoire constant, peu importe la nature de l'en-cours, présente un avantage certain: les flux seront plus réguliers. Ainsi en régime de croisière et après une période d'adaptation, le contrôle de la production sera beaucoup plus aisé. Et on pourra mieux prédire le taux de production du système.

Nécessairement, cette façon de déterminer la taille des lots a un inconvénient majeur: que fait-on si l'on doit produire en deçà du volume fixé? En fait, rien n'empêche de lancer des lots de taille moindre. L'important c'est qu'une bonne proportion de la production (de l'ordre de 70% par

exemple) soit gérée selon la normalisation établie. Face à une situation hors-norme, on a qu'à introduire des cartes de production temporaires dans le système.

La complexité de la méthode CONWIP est de gérer les différentes listes de commandes atelier. Cette gestion doit être effectuée par les personnes responsables du contrôle de la production. On peut accorder un statut prioritaire à des commandes atelier, en les mettant en tête de liste. De plus, on peut en tout temps insérer de nouvelles commandes.

La méthode CONWIP est plus générale que la méthode Kanban. En utilisant des cartes de production génériques, la demande n'a pas besoin d'être hautement répétitive. De plus, il n'est pas nécessaire de maintenir des en-cours pour chaque composant.

La méthode Kanban n'est pas appropriée pour des environnements de production qui ont des temps de mise en course élevés. Mais avec la méthode CONWIP, on peut minimiser ces temps lors de l'établissement des listes de commandes atelier.

Le jalonnement des ordres de fabrication à l'entrée d'une cellule est effectué par un personnel approprié. En plus

de l'avantage mentionné ci-haut, on peut éventuellement améliorer la productivité d'une cellule en chevauchant certaines opérations, contrairement à la méthode Kanban où le jalonnement est effectué automatiquement.

Gérer un système manufacturier de type atelier spécialisé en s'inspirant de la méthode CONWIP peut être plus avantageux que de gérer ce même système en flux poussés. Car en plus de profiter des techniques du groupement analogique, on profite également des avantages reliés au JAT, qui ont déjà été présentés.

Le concept de CONWIP est similaire au raisonnement de la gestion par les contraintes connue sous le vocable DBR ("Drum-Buffer-Rope"). DBR (voir Goldratt (1988)) est plus générale que CONWIP, car elle est applicable à un environnement de type atelier spécialisé dans sa forme la plus pure. Cependant, il existe plusieurs similarités entre ces deux méthodes si on base notre raisonnement sur les flux de production.

Sous une gestion DBR, la cadence de production est dictée par le goulot d'étranglement. Les opérations en amont de celui-ci sont programmées de telle sorte qu'elles maintiennent un volume d'en-cours suffisant en avant du goulot

d'étranglement, permettant ainsi une alimentation continue en composants. La corde sert à "tirer" la quantité juste de matière nécessaire.

Sous une gestion CONWIP, c'est également le goulot d'étranglement qui détermine la cadence de production. Les stocks en amont de ce dernier apparaissent naturellement, si le volume d'en-cours visé est suffisant. Enfin, c'est seulement lorsqu'une tâche est terminée que l'on peut en lancer une nouvelle (la corde).

La principale différence entre ces deux modes de gestion, c'est que sous une gestion DBR on estime le temps de passage nécessaire à un lot pour qu'il parvienne au goulot, tandis que sous une gestion CONWIP, les lancements sont automatiques.

CHAPITRE II

LA FORMULATION DU MODÈLE

Le système de production à l'étude

Le système de production étudié au cours de cette recherche est celui d'une entreprise qui oeuvre dans le domaine du meuble et bois ouvré. Ce système est constitué d'un atelier spécialisé comportant 37 postes de travail, atelier qui alimente 4 lignes d'assemblage. La Figure 1. présente schématiquement cette usine.

L'entreprise produit en lots intermittents pour stock de produits finis, dans un contexte où la demande est hautement variable et où les structures de produits varient considérablement, allant d'un niveau 0 jusqu'à un niveau 3.

Les hypothèses émises

Cette section dresse la liste des caractéristiques et conditions sous lesquelles le système a été modélisé. Modéliser exactement le système de production à l'étude est presque impossible à réaliser, cependant la modélisation utilisée offre une bonne approximation de celui-ci.

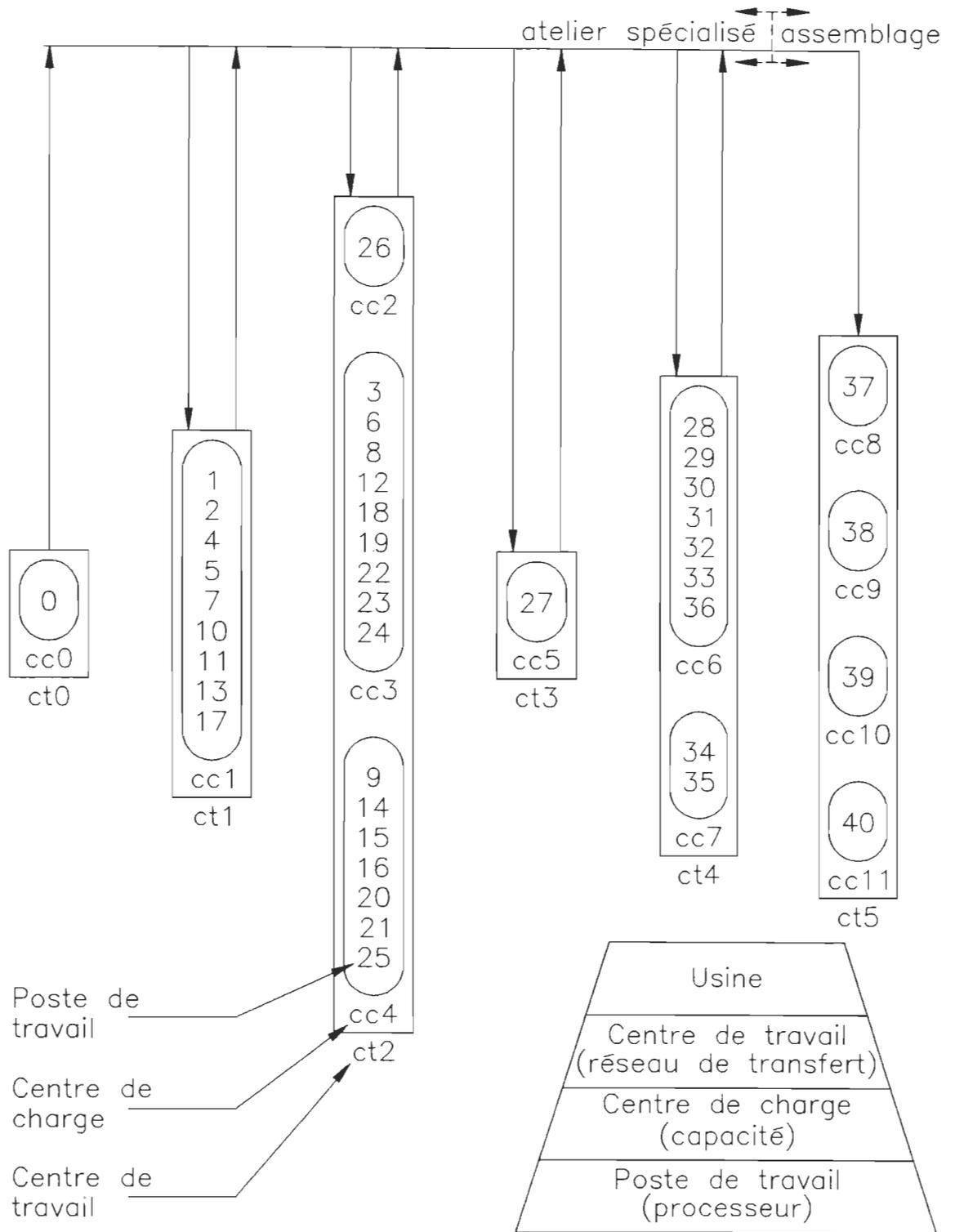


Figure 1. Schématique du système productif modélisé

Les hypothèses utilisées ont été classées selon qu'elles sont reliées: (a) à la capacité (ressources), (b) à la charge (tâches et caractéristiques), (c) aux procédures de contrôle et (d) aux critères de performance.

La capacité de l'usine

L'usine est constituée de 11 centres de charge, auxquels sont assignés des postes de travail différents ainsi que des ressources humaines. Chaque centre de charge a une capacité horaire journalière qui lui est propre et qui peut varier entre 8 heures et 18 heures de travail. Néanmoins, la capacité horaire journalière demeure la même au cours d'un horizon de planification donné (aucun temps supplémentaire). Le calendrier de production employé est constitué de semaines de cinq jours de travail et de mois de quatre semaines.

Les centres de charge de zéro à sept composent la partie atelier spécialisé de l'usine et lorsqu'un de ces centres n'est composé que d'un poste de travail, c'est que ce dernier correspond à un goulot d'étranglement qui limite continuellement la capacité du système. Les centres de charge ont été constitués de façon à regrouper les postes de travail qui sont souvent visités par les mêmes tâches, et les retours en arrière sont permis à l'intérieur de ceux-ci, mais lorsqu'une tâche sort d'un centre de charge, elle n'a plus à y

retourner. Les centres de charge de 8 à 11 correspondent aux 4 lignes d'assemblage.

Le niveau de main-d'oeuvre varie d'un centre à l'autre, mais il demeure constant pour un horizon de planification donné. On présume qu'un ouvrier travaille continuellement à 100% de rendement. Et un ouvrier ne peut jamais être transféré à un autre centre de charge.

Les autres hypothèses employées sont:

- (a) les postes de travail sont hautement fiables;
- (b) un poste de travail devient inactif uniquement lorsqu'on n'a pas de tâche ou d'ouvrier à lui affecter, et une tâche ne peut pas être retardée parce qu'il manque un outil spécial;
- (c) la capacité de l'usine est fixe, on ne considère pas l'absentéisme, l'indolence, etc...

Un centre de travail est constitué d'un ou plusieurs centres de charge. Ces centres de travail ont été constitués de façon à ce qu'il n'y ait pas de retour en arrière possible d'un centre de travail à un autre. Ceux-ci seront utilisés en flux tirés seulement afin de limiter, lorsque désiré, les transferts de lots d'un centre à l'autre.

Les tâches et leurs caractéristiques

Elvers (1974) a effectué plusieurs simulations afin de déterminer si les distributions de taux d'arrivée influençaient la performance relative de diverses règles de priorité utilisées pour ordonnancer les opérations. La génération des commandes était effectuée une fois par semaine. Les distributions avaient toutes la même moyenne. Et le système de production était sollicité à 79% de sa capacité. En utilisant seize distributions de taux d'arrivée réparties en six catégories, et en considérant dix règles de priorité, il est arrivé à la conclusion que les distributions n'étaient pas des variables significatives lorsque l'on compare l'efficacité relative de diverses règles de priorité.

Ce résultat est important, car il simplifie la méthodologie à appliquer pour générer les commandes. Dans le cas qui nous intéresse, nous effectuons également une génération périodique des commandes clients, en tenant compte

d'un horizon de planification de six semaines. Notons qu'une commande client peut être constituée d'un ou de plusieurs produits différents. La date d'arrivée d'une commande client dans le système correspond à la date à laquelle celle-ci est prise en compte par le planificateur. La date promise est affectée à l'aide d'une constante égale à quatre semaines. Il n'y a pas de fractionnement de commande à la livraison. La date de sortie du système correspondra à la date à laquelle une commande client est complétée.

L'entreprise fabrique une multitude de produits, mais plusieurs de ceux-ci sont similaires. Pour l'étude, nous avons utilisé 40 produits types. Un produit type étant un produit représentatif d'une famille de produits, ces produits types, dont chacun a sa propre structure et nomenclature, ont été employés afin de représenter les 252 produits gérés pour stocks de produits finis.

La détermination des besoins en produits finis (en termes de quantités et de dates promises) est effectuée à l'aide d'une gestion des stocks prévisionnelle. Par la suite, ceux-ci sont utilisés pour générer des ordres de fabrication de composants ou d'assemblages ayant des dates promises et de lancement qui tiennent compte d'un jalonnement à capacité infinie. Pour un jour donné, les commandes usine et les ordres

de fabrication arrivent simultanément au début de la période de travail.

Chaque opération possède des temps de mise en course et des temps opératoires. Ces durées varient selon des distributions uniformes, à l'aide de deux paramètres propres à chaque poste de travail.

Les autres hypothèses émises sont:

- (a) chaque ordre de fabrication possède un seul routage (pas de routage alternatif);
- (b) le temps d'opération à un poste de travail varie linéairement selon la taille du lot correspondant;
- (c) le temps d'opération à un poste de travail varie linéairement selon le nombre d'opérateurs affectés;
- (d) la charge est toujours égale ou inférieure à la capacité pour un horizon de planification donné et pour un centre de charge donné;

- (e) les lots de production sont égaux aux lots de transfert, et le temps de transfert est négligeable.

Les procédures de contrôle

Les procédures de contrôle se divisent en trois parties: (a) l'ordonnancement des tâches en attente devant un poste de travail; (b) l'affectation de travailleurs aux postes de travail; et (c) l'affectation de kanbans génériques afin de permettre aux tâches d'accéder à l'intérieur d'un centre de travail (sous une gestion en flux tirés uniquement).

L'ordonnancement des opérations est décentralisé. Car l'ordonnancement centralisé se base sur des temps opératoires et sur des capacités prévisionnelles sans tenir compte des aléas, les données de départ sont donc forcément fausses. De plus, pour obtenir de bons résultats en ordonnancement centralisé, il faut disposer d'un outil informatique sophistiqué et dispendieux, ce qui est à éviter pour le système de production à l'étude.

En ordonnancement décentralisé, les lots progressent dans le système selon la logique des gammes, lorsqu'un poste se libère, on sélectionne dans la file un lot en attente en fonction d'un indice de priorité. Généralement, il suffit

d'utiliser peu de règles de priorité pour obtenir de bons résultats par simulation. Panwalkar et Iskander (1977) ont recensé plus d'une centaine de ces règles. Cependant, plusieurs d'entre elles ne sont que des variantes de règles maîtresses. Même si ces règles sont simples d'utilisation, il n'en demeure pas moins qu'il est difficile d'établir à quel moment et à quel poste ou centre on doit utiliser telle ou telle règle. La meilleure façon de procéder est de simuler plusieurs scénarios. Par contre, on peut s'inspirer de la littérature pour converger rapidement vers de bons résultats.

Ramasesh (1990) a passé en revue les publications portant sur l'ordonnancement en atelier spécialisé et en se limitant aux travaux de recherche réalisés par simulation. Il présente plusieurs études qui démontrent que la nature des temps opératoires affecte significativement l'efficacité relative des règles de priorité. Par exemple, un temps opératoire distribué exponentiellement favorise la règle TOM ("Temps Opérateur Minimum") lorsque des temps opératoires longs sont générés. Pour contrer ce problème, on ne doit cependant pas se rabattre sur des temps déterministes, car ceci favoriserait une surévaluation de la performance. Ici, des distributions uniformes sont utilisées pour diverses raisons: (a) cette distribution est représentative de la plupart des temps de mise en course et opératoires observés; (b) ce type de

distribution est très facile à utiliser puisque l'on n'a pas besoin de la tronquer; et (c) cette distribution est plus facile à interpréter, ce qui facilite la collecte de données.

On sait également que la règle TOM semble être la règle la plus performante si l'on veut minimiser le temps d'achèvement moyen ou le temps d'attente moyen ainsi que le nombre de tâches sur le plancher de l'usine. De plus, cette règle semble plus robuste face aux variations des temps opératoires. Elle mérite ses lettres de noblesse surtout lorsque c'est un système fortement sollicité ou dans lequel les dates promises sont assez contraignantes. Par contre, elle augmente la variation des flux et elle peut retarder énormément des opérations. Lorsque c'est un système faiblement chargé ou dont les dates promises sont peu contraignantes, il est préférable d'utiliser des règles basées sur la marge.

Malgré la supériorité apparente de la règle TOM, il semble qu'elle soit peu efficace pour un environnement de production de type assemblage, selon Ramasesh (1990). Dans ce cas, de simples règles basées sur la convergence de composantes pour l'assemblage performant mieux.

Lorsqu'une opération est terminée à un poste de travail donné, le lot se déplace vers le prochain poste de

travail selon son routage. Si elle visite un nouveau centre de travail, elle sera en attente devant celui-ci avant d'accéder au prochain poste de travail. Si la prochaine opération est une opération d'assemblage, le lot est mis en attente de regroupement jusqu'à ce que les autres composants ou sous-assemblages le rejoignent. Si l'opération effectuée était la dernière, le lot est libéré et la position physique du stock visé par celui-ci est ajustée.

Lorsqu'un lot arrive dans la file d'attente d'un centre de travail, sa gamme est examinée afin de déceler si il doit ultérieurement visiter un prochain centre de travail. Si oui et si le transfert entre ces deux centres est limité par des kanbans génériques et qu'il n'y en a pas de disponible pour ce lien, il demeure en attente. Autrement, le lot peut accéder au poste de travail visé.

À ce moment, si on avait antérieurement affecté un kanban générique à ce lot, le kanban redevient disponible et retourne au centre de travail en amont correspondant. Si des lots sont en attente pour ce kanban, on affecte ce dernier au lot qui a la plus haute priorité (le calcul des priorités est effectué à ce moment, et non lorsqu'un lot arrive dans la file d'attente).

Lorsqu'un lot arrive dans la file d'attente d'un poste de travail, si ce poste est libre et qu'un nombre compris entre un minimum et un maximum de travailleurs (limites attribuées à chaque poste de travail) est disponible pour le centre de charge visé, on lui affecte le plus d'opérateurs possible et le lot est traité. Le poste de travail est alors occupé. Si un autre lot arrive dans la file d'attente de ce poste pendant que celui-ci est occupé, le lot demeure en attente.

Lorsque le nombre d'opérateurs disponibles est insuffisant pour traiter le lot, ce dernier demeure en attente devant le poste de travail. On affecte les opérateurs disponibles (s'il y a lieu) à ce poste et celui-ci sera en attente d'autres opérateurs. Les opérateurs ainsi affectés sont inactifs. Et le poste sera considéré indisponible jusqu'à ce qu'un nombre suffisant de travailleurs lui soit affecté.

Lorsque le traitement d'un lot est terminé, les travailleurs qui étaient utilisés redeviennent disponibles pour le centre de charge. Si un poste est en attente d'opérateurs et qu'on avait déjà débuté à lui affecter des opérateurs, on lui en affecte encore plus. Si on atteint un nombre suffisant d'opérateurs et qu'il y a encore des opérateurs disponibles et qu'aucun autre poste inactif n'est en attente de travailleurs

dans le centre de charge, on affectera le plus d'opérateurs possible. Ensuite, le poste de travail pourra traiter le lot le plus prioritaire dans sa file d'attente à ce moment.

S'il n'y a pas un poste en attente dont on avait débuté l'affectation de travailleurs, ou qu'il y en avait un mais on lui a affecté le minimum d'opérateurs nécessaires, qu'il y a encore des opérateurs disponibles et que plus d'un poste de travail est en attente d'opérateurs, on affectera les opérateurs disponibles dans un premier temps au poste ayant le lot le plus prioritaire, en considérant l'ensemble des lots dont leur poste de travail est en attente de travailleurs.

Finalement, s'il reste des travailleurs disponibles après avoir satisfait au moins au minimum les postes de travail qui étaient en attente d'opérateurs, on considère l'ensemble des postes actuellement occupés dont le nombre maximum d'opérateurs affectés n'a pas été atteint. Et l'on affectera au poste traitant la tâche la plus prioritaire le plus de travailleurs possibles. Et ainsi de suite jusqu'à ce qu'il n'y ait plus d'opérateurs disponibles, ou que l'on ne puisse pas affecter davantage d'opérateurs aux postes de travail.

Les autres hypothèses émises sont:

- (a) le traitement d'une tâche à un poste de travail peut être interrompu pour lui affecter plus de travailleurs, mais pas pour lui en enlever;
- (b) si à la fin d'un horizon de planification, le niveau de main-d'oeuvre varie à la baisse pour un centre de charge, on attendra que les opérateurs redeviennent disponibles avant d'en enlever (pas d'interruption pour les tâches en traitement);
- (c) si à la fin d'un horizon de planification, le niveau de main-d'oeuvre varie à la hausse pour un centre de charge, on suit immédiatement la séquence d'affectation des travailleurs disponibles décrite.

Les critères de performance

Les critères de performance employés pour une période de production donnée sont le temps de passage moyen des commandes usine et le niveau quotidien d'en-cours moyen. Le

temps de passage moyen est un critère très utilisé qui donne une bonne indication sur la réactivité d'un système de production. Et le niveau d'en-cours moyen observé à la fin de chaque journée sur le plancher de l'usine est une information très utile dans un environnement industriel (mouvement de trésorerie, aménagement d'usine, etc...).

Formulation mathématique

Les définitions et les expressions employées sont présentées dans le même ordre que celui des hypothèses.

La capacité de l'usine

Au début d'un horizon de planification h donné, on génère pour l'horizon de planification $h+1$ au complet l'arrivée quotidienne des commandes clients. Le nombre de commandes clients à piger pour une journée varie selon une distribution uniforme. Et la pige est effectuée à partir d'un bassin de commandes clients. Ce bassin contient la moitié des commandes clients de l'entreprise à l'étude sur une période d'une année. L'échantillonnage a été prélevé de façon systématique.

Au fur et à mesure que la génération est effectuée, on somme pour chaque produit la demande réelle pour l'horizon de planification visé. Les commandes générées demeurent

emmagasinées en mémoire jusqu'à ce que leur date d'arrivée concorde avec l'avancement de la simulation.

Sitôt la génération terminée pour un horizon de planification $h+1$, on calcule les quantités à fabriquer pour chaque produit à l'horizon h afin de satisfaire la demande à l'horizon $h+1$. On utilise une gestion des stocks selon une demande stochastique et une règle de décision basée sur le niveau de service recherché.

Pour chaque produit p géré sur stock, on emploie les variables et paramètres suivants arrondis à l'unité près:

$DR_{p,h}$ la demande réelle pour l'horizon h ;

PV la précision sur les prévisions exprimée en pourcentage;

v une valeur qui varie selon une distribution uniforme entre $-PV/100$ et $+PV/100$,
($|v| < 1$);

$DP_{p,h}$ la demande prévisionnelle pour l'horizon h , pour l'établir, il s'agit d'induire une erreur de prévision;

$$DP_{p,h} = (1+v)DR_{p,h} \quad (2,1)$$

NS le niveau de service recherché exprimé en pourcentage (le niveau de non-pénurie);

FS le facteur de sécurité;

$$FS = \left[\frac{0.2887}{0.2887 - (0.5 - (100 - NS)/100) \sqrt{(2PV/100)^2/12}} \right] \quad (2,2)$$

$Q_{p,h}$ la quantité totale à fabriquer pour l'horizon h , afin de satisfaire la demande pour l'horizon $h+1$;

$$Q_{p,h} = FS \cdot DP_{p,h+1} - FS \cdot DP_{p,h} + DR_{p,h} - SP_{p,h} \quad (2,3)$$

$SP_{p,h}$ le stock prévisionnel pour l'horizon h (la position de l'inventaire).

$$SP_{p,h+1} = FS \cdot DP_{p,h} - DR_{p,h} + SP_{p,h} + Q_{p,h} - FS \cdot DP_{p,h+1} \quad (2,4)$$

Un exemple d'utilisation du facteur de sécurité est présenté à l'Annexe B.

Par la suite et à l'aide des quantités à fabriquer obtenues pour l'horizon h , on calcule pour chaque produit le nombre de commandes usine à lancer et leur taille selon la quantité économique à commander spécifique à chaque produit sous une gestion en flux poussés, ou selon un fractionnement qui tient compte des goulots d'étranglement en flux tirés.

Sous une gestion en flux poussés et pour un produit donné, considérons les variables et paramètres suivants:

$Q_{p,h}$ la taille d'une commande usine à

l'horizon h ;

QEC la quantité économique à commander;

PQ la variation admissible sur la quantité économique à commander, exprimé en pourcentage.

On cherche à résoudre par itération le système d'équations suivant:

$$\text{MINIMISER } \sum_{1 \leq n \leq L}^n \left\{ Q_{p,h(n)} - QEC \right\}^2 \quad (2,5)$$

SUJET À

$$\sum_{l=1}^n Q_{p,h(n)} = Q_{p,h} \quad (2,6)$$

$$Q_{p,h(n)} \leq \left(1 + \frac{PQ}{100}\right) QEC \quad , n = 1, 2, \dots, L \quad (2,7)$$

Pour une gestion en flux tirés et pour un produit donné, considérons également les variables et paramètres suivants:

k les numéros de postes de travail qui limitent continuellement la capacité du système, ($k = 1, 2, \dots, K$);

T_k la durée cible d'un ensemble d'opérations au goulot k ;

$t_{p,ij,k}^{\min}$ la durée estimée pour effectuer l'opération i au goulot k de la composante j avec un minimum de travailleurs au poste de travail;

$q_{p,j}$ la quantité de composants de type j nécessaires pour fabriquer une unité du produit;

$$t_{p,i,k} = \sum_{j=1}^J q_{p,j} \cdot t_{p,ij,k}^{\min} \quad (2,8)$$

$s_{p,ij,k}^{\min}$ la durée estimée pour effectuer la mise en course i au goulot k de la composante j avec un minimum de travailleurs au poste de travail;

$$s_{p,i,k} = \sum_{j=1}^J s_{p,ij,k}^{\min} \quad (2,9)$$

$S_{p,k}$ l'ensemble des composants d'un produit qui visitent le goulot k .

$$C_{p,k} = \sum_{i \in S_{p,k}} t_{p,i,k} + s_{p,i,k} \quad (2,10)$$

$$k^b = \operatorname{argmax}_k C_{p,k} \quad (2,11)$$

On cherche à résoudre par itération le système d'équations suivant:

$$\operatorname{MINIMISER}_{1 \leq n \leq L} \sum_{l=1}^n \left\{ \sum_{i \in S_{p,k^b}} (Q_{p,h(n)}^l \cdot t_{p,i,k^b} + n \cdot s_{p,i,k^b}) - T_k \right\}^2 \quad (2,12)$$

SUJET À

$$\sum_{l=1}^n Q_{p,h(n)}^l = Q_{p,h} \quad (2,13)$$

$$\sum_{i \in S_{p,k^b}} (Q_{p,h(n)}^l \cdot t_{p,i,k^b} + n \cdot s_{p,i,k^b}) \leq (1 + \frac{PQ}{100}) T_k, \quad n = 1, 2, \dots, L \quad (2,14)$$

Après avoir déterminé pour un produit le nombre de commandes usine à lancer et leur taille à l'horizon h pour satisfaire la demande à l'horizon $h+1$, on affecte à chaque commande une date promise afin de répartir la production.

Soit les variables suivantes tronquées à l'unité près:

d_h^0 la date au début de l'horizon h;

$dp_{p,h(n)}$ la date promise de la n^e commande usine à l'horizon h pour le produit p.

Les dates promises pour chaque commande usine d'un produit donné à l'horizon h sont:

$$dp_{p,h(n)} = d_h^0 + n \cdot \frac{(d_{h+1}^0 - d_h^0)}{L} \quad , \quad n = 1, 2, \dots, L$$

$$, \quad \frac{(d_{h+1}^0 - d_h^0)}{L} \geq 1 \quad (2,15)$$

Ensuite, on calcule pour chaque centre de charge cc (cc = 0,1,...,11) sa capacité horaire journalière ainsi que son niveau de main-d'oeuvre pour l'horizon de planification. Soit:

w un numéro de poste de travail (w = 0,1,...,40);

FMH un facteur de majoration pour la capacité horaire journalière compris entre zéro et un;

NT_w^{\min} le nombre minimum de travailleurs nécessaires

pour effectuer une opération au poste w;

NT_{cc}^{\min} le nombre minimum de travailleurs à affecter au centre de charge cc;

$$NT_{cc}^{\min} = \text{MAX}_{w \in cc} (NT_w^{\min}) \quad , \quad cc = 0, 1, \dots, 11 \quad (2,16)$$

NT_w^{\max} le nombre maximum de travailleurs pour effectuer une opération au poste w;

NT_{CC}^{\max} le nombre maximum de travailleurs à affecter au centre de charge cc ;

$$NT_{CC}^{\max} = \text{MIN}_{w \in cc} (NT_w^{\max}), cc = 0, 1, \dots, 11 \quad (2,17)$$

$S_{p,cc}$ l'ensemble des opérations d'un produit effectuées au centre de charge cc ;

$TW_{h,cc}$ le travail total estimé au centre de charge cc pour un horizon h .

$$TW_{h,cc} = \sum_{p=1}^{252} Q_{p,h} \cdot t_{p,i} + L_p \cdot S_{p,i}, cc = 0, 1, \dots, 11$$

$$, i \in S_{p,cc} \quad (2,18)$$

On cherche à résoudre le système d'équations suivant:

$$\text{MINIMISER}_{0 \leq cc \leq 11} (Y_{z,cc} + X_{z,cc}) \cdot NT_{CC} \quad (2,19)$$

SUJET À

$$NT_{CC} \geq NT_{CC}^{\min}, cc = 0, 1, \dots, 11 \quad (2,20)$$

$$NT_{CC} \leq NT_{CC}^{\max}, cc = 0, 1, \dots, 11 \quad (2,21)$$

$$X_{z,cc} (8.00 + 0.50z) \geq \frac{TW_{h,cc}}{\text{FMH}(NT_{CC}^{\min} / NT_{CC}^{\max}) (d_{h+1}^o - d_h^o)}, z = 0, 1, \dots, 20$$

$$, cc = 0, 1, \dots, 11 \quad (2,22)$$

$$Y_{z,cc} = 0.25(z-8), z = 1, 2, \dots, 15$$

$$, cc = 0, 1, \dots, 11 \quad (2,23)$$

$$Y_{z,cc} = 0.25(z-16) \quad , z = 17, 18, \dots, 20$$

$$, cc = 0, 1, \dots, 11 \quad (2,24)$$

$$Y_{z,cc} = 0 \quad , z = 0, 16$$

$$, cc = 0, 1, \dots, 11 \quad (2,25)$$

$$\sum_{z=0}^{20} X_{z,cc} = 1 \quad , cc = 0, 1, \dots, 11 \quad (2,26)$$

$$X_{z,cc} = 0 \text{ ou } 1 \quad , \forall z, cc \quad (2,27)$$

$$Y_{z,cc} \text{ un nombre réel} \quad , \forall z, cc \quad (2,28)$$

À l'équation (2,19), on cherche à minimiser la rémunération des travailleurs dont la rémunération du temps supplémentaire est représentée par les équations (2,23), (2,24) et (2,25). On obtient une capacité horaire journalière à la demi-heure près qui est fonction du nombre d'opérateurs affectés.

Sous une gestion en flux tirés, on affecte un nombre de kanbans génériques, par horizon de planification, pour les liens entre les centres de travail 0-3, 1-3, et 2-3 à l'aide des coefficients α , β , et γ respectivement. Ces coefficients, compris entre zéro et un, demeurent constants lors d'une simulation. Si on désigne par $NTR_h, (\alpha, \beta, \text{ ou } \gamma)$ le nombre de transferts comptabilisés pour les trois liens pré-cités et pour un horizon h donné, et si l'on utilise $NK_h, (\alpha, \beta, \text{ ou } \gamma)$ comme étant le nombre de kanbans génériques entre ces liens pour le même horizon, alors ce dernier sera égal à:

$$NK_h, (\alpha, \beta, \text{ ou } \gamma) = (\alpha, \beta \text{ ou } \gamma) \frac{NTR_h, (\alpha, \beta, \text{ ou } \gamma)}{\begin{pmatrix} o & o \\ (d_{h+1} & -d_h) \end{pmatrix}} \quad (2,29)$$

Les tâches et leurs caractéristiques

À l'aide des commandes usine et de leur date promise, on affecte des dates de lancement et des dates promises aux ordres de fabrication. Soit:

$tpas_1$	le temps de passage estimé des ordres dans l'usine qui ne visitent aucun goulot;
$tpas_2$	le temps de passage estimé des ordres dans l'usine qui visitent le goulot correspondant au centre de charge 0;
$tpas_3$	le temps de passage estimé des ordres dans l'usine qui visitent les goulots aux centres de charge 0 et 5 (si une tâche visite le centre de charge 5, elle visite obligatoirement le centre de charge 0);
$tpas_x$	le temps de passage estimé d'un ordre selon son routage ($x = 1, 2$ ou 3);
$tpas_{CU}$	le temps de passage estimé des commandes usine, ce temps servira de borne supérieure;
$dl_{p,h(n)}^{j,niv}$	la date de lancement de l'ordre visant à fabriquer la composante j au niveau niv du produit p à l'horizon h pour la commande usine n ;
$dp_{p,h(n)}^{j,niv}$	la date promise de l'ordre visant à fabriquer la composante j au niveau niv du produit p à l'horizon h pour la commande usine n .

Les dates promises et de lancement (la date de lancement est temporaire) au niveau 0 sont:

$$dp_{p,h(n)}^{p,0} = dp_{p,h(n)} \quad (2,30)$$

$$dl_{p,h(n)}^{p,0'} = dp_{p,h(n)} - t_{pas_x} \quad (2,31)$$

Les dates promises et les dates temporaires de lancement au niveau 1 sont:

$$dp_{p,h(n)}^{j,1} = dl_{p,h(n)}^{p,0'} \quad , \forall j \in p \text{ au niveau 1} \quad (2,32)$$

$$dl_{p,h(n)}^{j,1'} = dp_{p,h(n)}^{j,1} - t_{pas_x} \quad , \forall j \in p \text{ au niveau 1} \quad (2,33)$$

Les dates promises et les dates temporaires de lancement au niveau 2 sont (il en va de même pour le niveau 3):

$$dp_{p,h(n)}^{j,2} = dl_{p,h(n)}^{j,1'} \quad , \forall j \text{ au niveau 2} \in j \text{ au niveau 1} \quad (2,34)$$

$$dl_{p,h(n)}^{j,2'} = dp_{p,h(n)}^{j,2} - t_{pas_x} \quad , \forall j \in p \text{ au niveau 2} \quad (2,35)$$

Il ne reste plus qu'à déterminer les dates de lancement finales pour tous les ordres de fabrication:

$$dl_{p,h(n)}^{j,niv} = \text{MAX} \left\{ dp_{p,h(n)} - t_{pas_{CU}}, \text{MIN} \left[dl_{p,h(n)}^{j,niv'} \right] \right\} \quad , \forall j \in p \quad (2,36)$$

Cette date correspond également à la date d'entrée $r(CU)$ de la commande usine.

Ainsi, les ordres de fabrication pour une commande usine donnée ont la même date de lancement. Car le poste de travail 0 (centre de charge 0) est une scie servant à débiter des panneaux, et que le débitage des composants d'une commande usine s'effectue en même temps. De cette façon, les patrons de coupe sont déterminés par commande usine.

Comme il a déjà été mentionné, la taille d'un lot de transfert est égale à la taille du lot à produire. Pour les lignes d'assemblage (postes de travail 37 à 40), cette hypothèse ne peut tenir car elle augmente indûment le temps de passage à l'assemblage. Pour contourner ce problème, les opérations d'assemblage ont été unifiées pour un ordre donné afin d'obtenir un seul temps opératoire et un seul temps de mise en course, et ce en considérant que les lots de transfert étaient de taille unitaire.

Pour un ordre donné qui doit être traité sur une ligne d'assemblage composée de W postes de travail, la mise en course unifiée est:

$$s_{p,ij,w}^{\min'} = \sum t_{p,ij,w}^{\min} + \text{MAX}\left\{s_{p,ij,w'}^{\min}\right\}, w' = 1, 2, \dots, W \quad (2,37)$$

Et son temps opératoire unifié devient:

$$t_{p,ij,w}^{\min'} = \text{MAX}\left\{t_{p,ij,w'}^{\min}\right\}, w' = 1, 2, \dots, W \quad (2,38)$$

Notons qu'à l'équation (2,37), on tient compte du temps nécessaire pour remplir la ligne d'assemblage (ou le temps pour la vider).

Les procédures de contrôle

Les procédures de contrôle employées dans cette recherche sont présentées au Tableau 1. Elles sont utilisées pour l'ordonnancement des ordres de fabrication en attente devant un poste de travail, pour l'affectation de travailleurs aux postes de travail et pour l'affectation de kanbans génériques pour les ordres en attente devant un centre de travail sous une gestion flux tirés.

Tableau 1

Liste et description des règles d'ordonnancement utilisées

Symbole	Description [†]
FIFO	Premier-arrivé, premier-servi
DLOFAMCU	Date de lancement de l'ordre de fabrication la plus ancienne, et après ce choix, traiter tous les ordres en attente appartenant à la même commande usine l'un après l'autre
DPOFP	Date promise de l'ordre de fabrication la plus proche
DPCUP	Date promise de la commande usine la plus proche
DPCUPMCU	Date promise de la commande usine la plus proche, et après ce choix, traiter tous les ordres en attente appartenant à la même commande usine l'un après l'autre
DPOFPMCU	Date promise de l'ordre de fabrication la plus proche, et après ce choix, traiter tous les ordres en attente appartenant à la même commande usine l'un après l'autre

[†] Si deux tâches ou plus ont le même indice de priorité, la règle utilisée est FIFO.

Les critères de performance

Si NCU commandes usine ont été complétées pour une période de production donnée, et si l'on note par C(CU) la date où une commande usine est complétée, le temps de passage moyen des commandes usine est égale à:

$$\bar{F} = \frac{\sum_{CU=1}^{NCU} (C(CU) - r(CU))}{NCU} \quad (2,39)$$

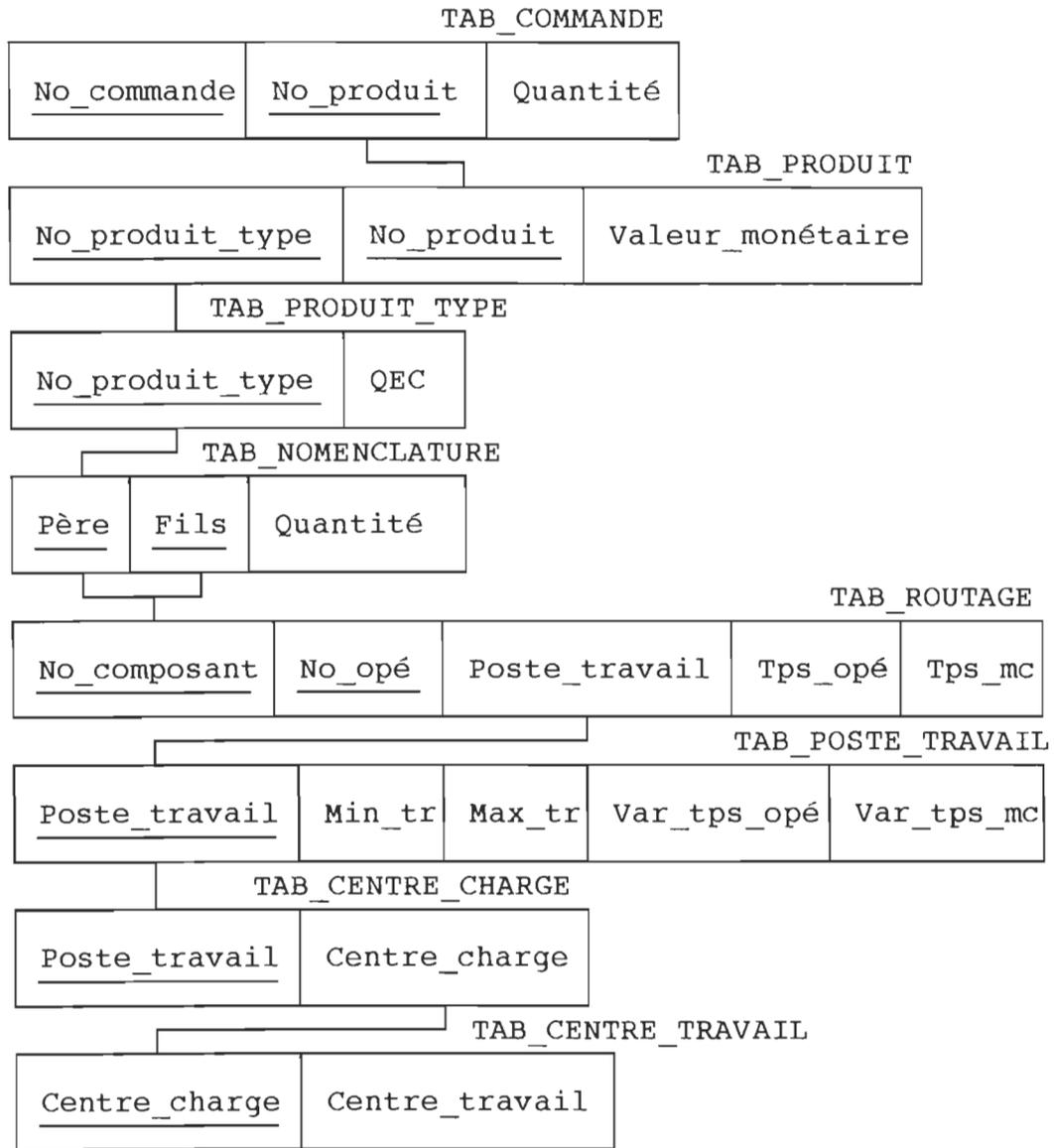
Le niveau quotidien d'en-cours est calculé au début d'une journée en sommant la taille de tous les lots en attente ou en cours de traitement dans l'usine, à l'exception des lots en attente devant le centre de travail 0 ou devant le centre de charge 0 (les composants n'ont pas encore été débités). Pour obtenir le niveau quotidien d'en-cours moyen, il suffit de diviser la somme des niveaux quotidiens d'en-cours par le nombre de jours correspondant à la période de production observée.

Pour calculer le nombre quotidien d'en-cours moyen associé aux chemins critiques, on considère uniquement les lots en attente devant les centres de travail 1, 2 et 3 ainsi que les lots à l'intérieur de ceux-ci. On exclut les lots en attente devant le centre de travail 0 ou devant le centre de charge 0.

Le design expérimental

L'usine et les tâches

Les données reliées à l'usine et aux tâches qui ont été utilisées dans cette recherche sont regroupées dans l'ouvrage de Poirier (1995). La Figure 2. montre schématiquement les tables de données concernées. La table TAB_COMMANDE contient le bassin des commandes clients pour un total de 4466 commandes. La table TAB_PRODUIIT contient les 252 produits gérés sur stocks. La table TAB_PRODUIIT_TYPE contient les 40 produits types ainsi que la taille de leur lot économique respectif.



 CHAMP CLEF

Figure 2. Les tables de données reliées au système de production à l'étude

La table TAB_NOMENCLATURE contient les relations père-fils-quantité utilisées pour la gestion de type MRP. La table TAB_ROUTAGE contient tous les routages pour les composants. Pour un composant donné, ce dernier possède des numéros d'opération chronologiques, des temps opératoires et des temps de mise en course. La table TAB_POSTE_TRAVAIL contient toutes les caractéristiques associées aux postes de travail. Les tables TAB_CENTRE_CHARGE et TAB_CENTRE_TRAVAIL contiennent les données nécessaires pour reproduire ce qui est présenté à la Figure 1.

Les paramètres expérimentaux

Le Tableau 2 présente les paramètres expérimentaux qui sont demeurés constants pour toutes les simulations. Ces paramètres ont été fixés de façon à se rapprocher le plus possible du système de production réel à l'étude.

Étant donné la multitude de variables expérimentales indépendantes à utiliser lors des simulations, les méthodes Taguchi ont été employées afin de minimiser le nombre de simulations. Le Tableau 3 montre la matrice orthogonale $L_{16}(2^{15})$ utilisée. Ainsi, 16 essais ont été effectués sous une gestion en flux poussés et tirés. Et chaque essai a été répété trois fois, pour un total de 96 simulations. Les semences principales utilisées lors de l'initialisation des générateurs de nombres pseudo-aléatoires pouvaient varier entre 0 et 32767. Elles ont été choisies à l'aide d'une table de nombres aléatoires.

Tableau 2

Les paramètres expérimentaux utilisés pour toutes les simulations

Symbole	Valeur	Description
n/a	38 à 46 unités	Le nombre quotidien de commandes clients à piger
PV	20 %	La précision sur les prévisions
PQ	20 %	La variation admissible sur la taille des commandes usine
FMH	20 %	Le facteur de majoration sur la capacité horaire journalière
n/a	4 semaines	La constante appliquée à la date d'arrivée d'une commande client afin de déterminer sa date promise
h	6 semaines	La durée d'un horizon de planification
tpas ₁	1 jour	Le temps de passage estimé d'un composant qui ne visite aucun goulot
tpas ₂	2 jours	Le temps de passage estimé d'un composant qui visite le goulot 0 (centre de charge 0)
tpas ₃	3 jours	Le temps de passage estimé d'un composant qui visite les goulots 0 et 1 (centres de charge 0 et 5)

Tableau 3

Méthodes TAGUCHI: table orthogonale $L_{16}(2^{15})$

Numéro d'essai	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
3	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2
4	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1
5	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2
6	1	2	2	1	1	2	2	2	2	1	1	2	2	1	1
7	1	2	2	2	2	1	1	1	1	2	2	2	2	1	1
8	1	2	2	2	2	1	1	2	2	1	1	1	1	2	2
9	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
10	2	1	2	1	2	1	2	2	1	2	1	2	1	2	1
11	2	1	2	2	1	2	1	1	2	1	2	2	1	2	1
12	2	1	2	2	1	2	1	2	1	2	1	1	2	1	2
13	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1
14	2	2	1	1	2	2	1	2	1	1	2	2	1	1	2
15	2	2	1	2	1	1	2	1	2	2	1	2	1	1	2
16	2	2	1	2	1	1	2	2	1	1	2	1	2	2	1

Le Tableau 4 montre les facteurs (variables expérimentales indépendantes) utilisés en flux poussés. La durée cible T_k d'un ensemble d'opérations au goulot k peut être considérée comme la durée cumulée moyenne recherchée pour exécuter une commande usine au goulot 0 ou 1. Cette valeur est employée dans le système d'équations (2,12), (2,13) et (2,14). Pour cette étude, on a utilisé cette façon de déterminer la taille des commandes usine, même en flux poussés, au lieu d'utiliser le système d'équations (2,5), (2,6) et (2,7). Ainsi, on élimine le biais qui aurait pu être induit à cause de la

détermination des quantités économiques à commander qui ont été fournies par l'entreprise participante.

Le coefficient sur le temps de mise en course C au goulot 0 ou 1 (selon que le goulot 0 ou 1 a été utilisé pour la détermination de la taille des commandes usine) comme facteur expérimental sert à déterminer si les temps de mise en course généralement observés pour ces goulots sont significativement longs. C'est-à-dire qu'ils peuvent affecter négativement la performance du système à l'étude.

Tableau 4

Facteurs expérimentaux en flux poussés

Numéro de colonne	Facteur
1	Durée cible d'un ensemble d'opérations au goulot k (T_k)
2	Coefficient sur le temps de mise en course au goulot 0 ou 1 (C)
3	Interaction ($T_k C$)
4	Temps de passage estimé des commandes usine (tp_{ascU})
5	Niveau de service recherché (NS)
6	Erreur résiduelle
7	Erreur résiduelle
8	Ensemble de règles d'ordonnancement (O)
9	Interaction ($T_k O$)
10	Interaction (CO)
11	Erreur résiduelle
12	Erreur résiduelle
13	Erreur résiduelle
14	Erreur résiduelle
15	Erreur résiduelle

L'emploi du temps de passage estimé des commandes usine $tpas_{CU}$ comme facteur permettra d'établir si la performance du système est sensible à une variation de cette borne supérieure utilisée dans l'équation (2,36). Avec le niveau de service recherché NS comme facteur, on vérifie si ce dernier peut influencer la performance du système de production. Finalement, un ensemble de règles d'ordonnement 0 est utilisé afin de déterminer si ce genre de système est très sensible au choix de règles d'ordonnement locales.

Le Tableau 5 présente les facteurs expérimentaux employés en flux tirés. Outre ceux pré-cités, on utilise également les coefficients α , β et γ que l'on retrouve dans l'équation (2,29).

Tableau 5

Facteurs expérimentaux en flux tirés

Numéro de colonne	Facteur
1	Durée cible d'un ensemble d'opérations au goulot k (T_k)
2	Coefficient sur le temps de mise en course au goulot 0 ou 1 (C)
3	Interaction ($T_k C$)
4	Coefficient pour la relation intercentre de travail 0-3 (α)
5	Coefficient pour la relation intercentre de travail 1-3 (β)
6	Temps de passage estimé des commandes usine ($tpas_{CU}$)
7	Erreur résiduelle
8	Ensemble de règles d'ordonnement (0)

Tableau 5 (suite)

Numéro de colonne	Facteur
9	Interaction ($T_k O$)
10	Interaction (CO)
11	Coefficient pour la relation intercentre de travail 0-3 (γ)
12	Niveau de service recherché (NS)
13	Erreur résiduelle
14	Erreur résiduelle
15	Erreur résiduelle

Les Tableaux 6 et 7 présentent les niveaux de ces facteurs pour les simulations en flux poussés et tirés, et ce en conformité avec le Tableau 3. Le facteur T_k est exprimé en heures, $tpas_{CU}$ en jours, NS, α , β , et γ en pourcentage.

Tableau 6

Assignation systématique des niveaux pour les facteurs en flux poussés

Numéro d'essai	T_k	C	O	$tpas_{CU}$	NS
1	α	1,0	1	6	50
2	α	1,0	2	6	50
3	α	1,0	1	7	95
4	α	1,0	2	7	95
5	α	0,5	1	6	50
6	α	0,5	2	6	50
7	α	0,5	1	7	95
8	α	0,5	2	7	95
9	3,5	1,0	1	6	95

Tableau 6 (suite)

Numéro d'essai	T_k	C	O	tpascu	NS
10	3,5	1,0	2	6	95
11	3,5	1,0	1	7	50
12	3,5	1,0	2	7	50
13	3,5	0,5	1	6	95
14	3,5	0,5	2	6	95
15	3,5	0,5	1	7	50
16	3,5	0,5	2	7	50

Tableau 7

Assignation systématique des niveaux pour les facteurs
en flux tirés

Numéro d'essai	T_k	C	O	tpascu	NS	α	β	γ
1	α	1,0	1	6	50	20	25	20
2	α	1,0	2	6	95	20	25	30
3	α	1,0	1	7	95	30	35	20
4	α	1,0	2	7	50	30	35	30
5	α	0,5	1	7	50	20	25	30
6	α	0,5	2	7	95	20	25	20
7	α	0,5	1	6	95	30	35	30
8	α	0,5	2	6	50	30	35	20
9	3,5	1,0	1	6	50	20	35	30
10	3,5	1,0	2	6	95	20	35	20
11	3,5	1,0	1	7	95	30	25	30
12	3,5	1,0	2	7	50	30	25	20
13	3,5	0,5	1	7	50	20	35	20
14	3,5	0,5	2	7	95	20	35	30
15	3,5	0,5	1	6	95	30	25	20
16	3,5	0,5	2	6	50	30	25	30

Enfin, le Tableau 8 montre les règles d'ordonnement locales employées aux niveaux un et deux. Notons que celles-ci ont été établies de façon à ce qu'on utilise pour un centre de travail donné la même règle aux centres de charge et aux postes de travail.

Tableau 8

Assignation des règles d'ordonnement locales

Poste de travail	Niveau un	Niveau deux
0	DLOFAMCU	DLOFAMCU
1 à 26	DPOFP	DPCUP
27	DPOFPMCU	DPCUPMCU
28 à 36	DPOFP	DPCUP
37 à 40	DPOFP	DPCUP

CHAPITRE III

LE SIMULATEUR ET LA VALIDATION DU MODÈLE

Le simulateur

Le progiciel de simulation a été développé dans un environnement WINDOWS à l'aide d'un micro-ordinateur possédant un processeur de type 80486. Ce micro-ordinateur était doté d'une mémoire étendue de 16 Mo, d'une mémoire virtuelle de 16 Mo ainsi que d'une mémoire morte de 40 Mo.

Le programme principal a été programmé en langage C++. Après l'initialisation, ce dernier appelait le réseau représentant l'usine modélisée en langage SLAM II. Au cours d'une simulation, c'est le réseau qui vivait virtuellement les différents événements, en appelant continuellement les procédures programmées en C++. D'une manière imagée, on peut dire que la partie "bureau" où s'effectue toute la gestion était modélisée en langage C++, et que le langage SLAM II était réservé à la partie "usine". La Figure 3. montre une vue globale du progiciel.

Comme le design du progiciel de simulation n'est pas le but premier de cette recherche, une brève description du réseau

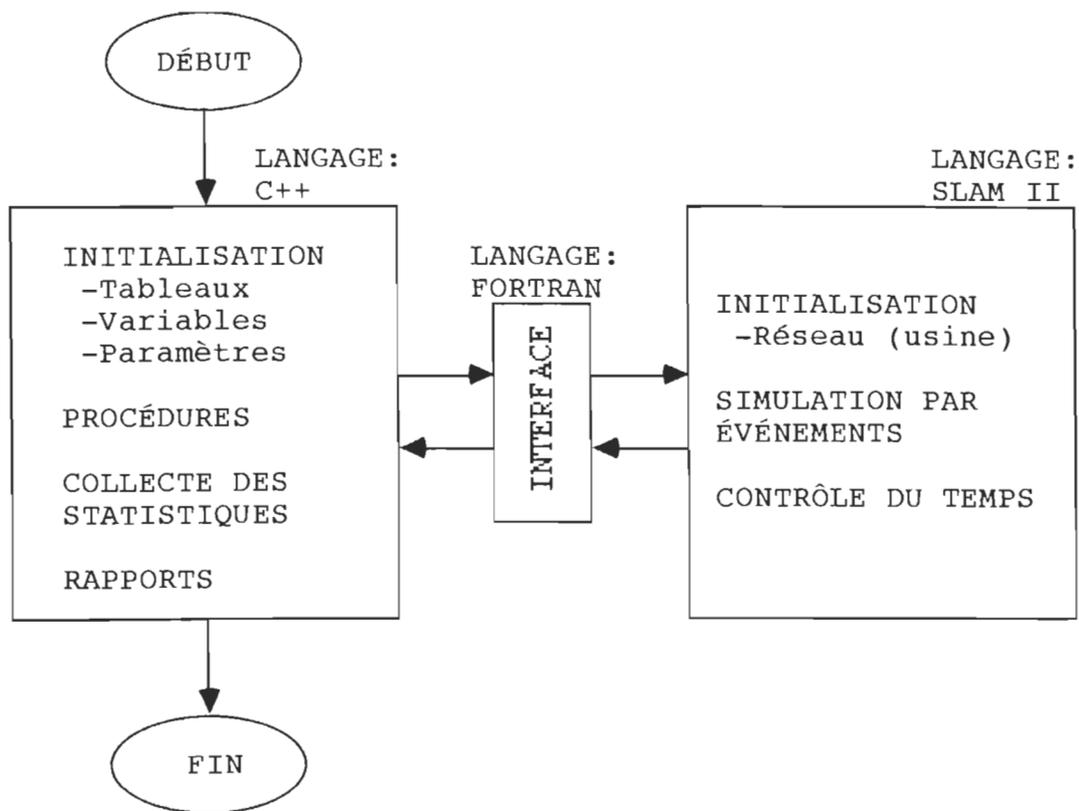


Figure 3. Vue globale du progiciel de simulation

en Langage SLAM II est présentée à l'Annexe A.

La durée d'une simulation variait en moyenne de 30 à 45 minutes selon que l'on simulait une gestion en flux poussés ou en flux tirés respectivement.

Le régime transitoire

Au début d'une simulation, toutes les ressources étaient disponibles, ce qui n'est pas représentatif de la réalité. Ainsi, la collecte des statistiques débutait après une période de 12 semaines.

La génération des commandes clients débutait au deuxième horizon de planification, soit pour les semaines de 6 à 12. Comme l'on simulait une production pour stocks, les commandes usine commençaient à être traitées lors du premier horizon. C'est au cours du deuxième horizon que le régime permanent s'installait. Puisque la durée du régime transitoire peut varier selon les niveaux des différents facteurs employés, la collecte ne débutait qu'au début de la treizième semaine.

Pour illustrer le régime transitoire, la Figure 4. présente l'évolution hebdomadaire du temps de passage moyen des commandes usine. Le temps de passage d'une commande usine étant collecté dans la semaine où cette commande est complétée. Les données utilisées proviennent d'une simulation sous une gestion en flux poussés et dont la charge de l'usine était de 80 %.

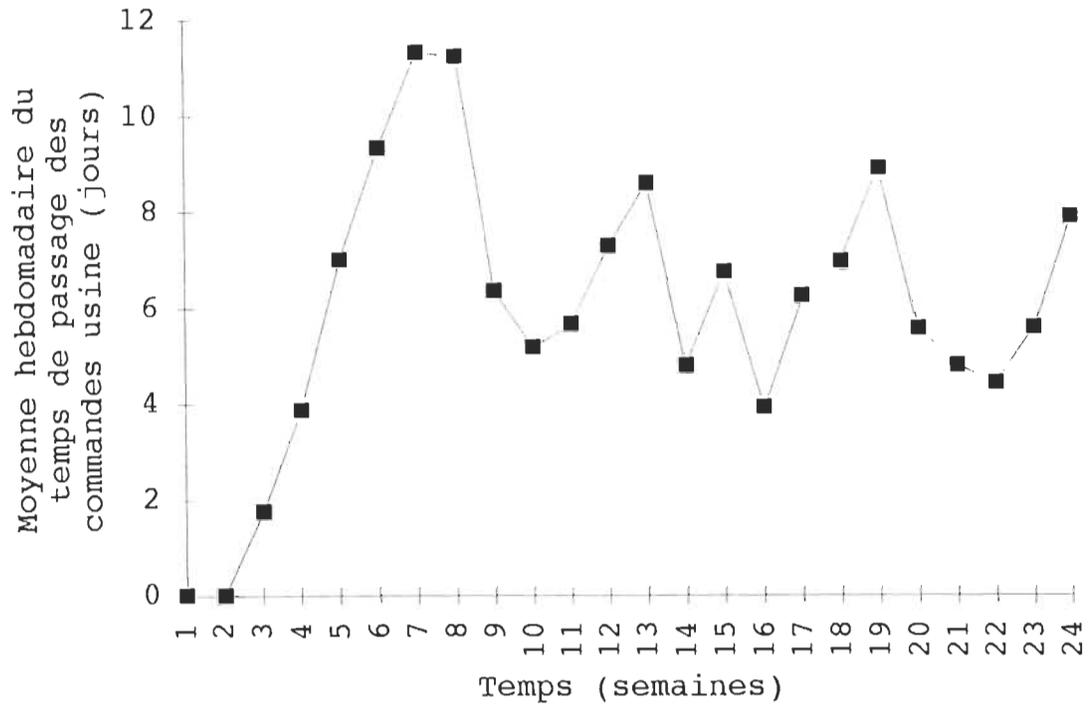


Figure 4. Transition entre les régimes transitoire et permanent observée pour une simulation donnée.

Le régime permanent

Le régime permanent, tel que discuté ici n'est pas un état d'équilibre pur. En fait, le temps de passage moyen des commandes usine et les autres mesures employées peuvent fluctuer pendant ce régime. Nécessairement, plus celui-ci est long et plus les mesures collectées tendront vers des valeurs à l'équilibre.

Considérant le temps machine fort coûteux pour une simulation, et considérant qu'il s'agit d'une étude comparative, la longueur des simulations a été établie selon l'objectif d'obtenir des valeurs de $m \pm s/4$, à un intervalle de confiance de 95 %. Comme l'on ne connaît ni σ , ni l'étendue des populations, on peut utiliser le théorème de Tchebycheff qui nous indique que la taille de l'échantillon devrait être d'au moins 320 valeurs collectées.

Cette quantité de commandes usine est atteinte au cours du sixième horizon de planification, pour l'exemple employé à la Figure 4. Mais si on considère le caractère aléatoire de la génération de la demande, il est préférable de terminer la simulation au cours du septième horizon. Et puisque la collecte débute au début d'un horizon, il est normal de la terminer à la dernière semaine d'un horizon, soit à la fin de la 42^e semaine.

La validation

La validation consiste à s'assurer d'une part que le simulateur effectue adéquatement ce dont on attend de lui, et d'autre part à valider le fonctionnement de celui-ci par rapport au système réel de production à l'étude.

Pendant le développement du progiciel, tous les modules et/ou procédures ont été vérifiés avec des valeurs représentatives de la réalité ainsi qu'avec des valeurs aux limites, afin de s'assurer du bon fonctionnement du simulateur.

Ensuite, des simulations préliminaires ont été effectuées en choisissant au hasard des essais inclus dans le design expérimental, sous une gestion en flux poussés, afin de comparer ces résultats avec le système réel de production. Notons que toutes les données d'entrée utilisées proviennent de l'entreprise et que la collecte de celles-ci s'est échelonnée sur plus d'un an et demi.

La capacité de l'usine

L'usine en question fixe la capacité de ses ressources sur un horizon de quatre à huit semaines. Ainsi, en utilisant un horizon de six semaines, on se rapproche de leur tactique de production. La main-d'oeuvre est flexible à l'intérieur d'un même département, il en est ainsi pour les centres de charge modélisés. Cependant, les centres de charge n'ont pas été nécessairement constitués de la même façon que les départements réels.

Les autres hypothèses employées se rapprochent de la réalité. Le modèle ne considère pas les pannes car pour le

système à l'étude, les pannes sont relativement rares et les temps de réparation sont courts. De plus, l'absentéisme chronique et l'indolence n'affectent qu'une infime partie des ouvriers de cette usine. Enfin, le temps supplémentaire non planifié est maintenu en deçà de 5 % de toutes les heures-hommes travaillées.

Notons que toutes les simulations ont été effectuées en sollicitant le système à 80 % de sa capacité. Il aurait certes été intéressant d'effectuer d'autres simulations avec des sollicitations plus élevées, mais la mémoire nécessaire sous SLAM II était déjà utilisée à son maximum.

Les tâches et leurs caractéristiques

Le niveau de génération des commandes clients employé engendrait des chiffres d'affaires approximatifs de 650 000\$ par horizon de planification. Ce qui est représentatif de la réalité lorsque l'usine produit avec un quart de travail.

Avec une telle sollicitation, le temps de passage moyen des commandes usine variait régulièrement entre 6 et 7 jours ouvrables, ce qui est encore représentatif de la réalité.

L'hypothèse que les temps de transfert sont négligeables n'a pas d'impact significatif, car ces temps sont

relativement courts et ils sont inclus dans les temps opératoires. Quant aux lots de production, ils ne sont pas toujours égaux aux lots de transfert, mais cette situation arrive rarement. Enfin, les autres hypothèses utilisées ont été posées à la suite d'observations effectuées à l'usine même.

Les procédures de contrôle

La plupart des hypothèses concernant les procédures de contrôle sont valides pour cette entreprise. Néanmoins, même si celles-ci demeurent constantes pour toute la durée d'une simulation, ce n'est pas toujours le cas en réalité. Par exemple, la modélisation ne tient pas compte de l'introduction dans le système d'une commande urgente possédant un haut niveau de priorité. Mais ce genre de situation est exceptionnel.

Les mesures de la performance

Les mesures de la performance employées sont considérées comme importantes par l'entreprise étudiée. Conséquemment, ils devraient l'être également pour d'autres entreprises du même secteur industriel.

Simuler exactement le système de production à l'étude est impossible. Il a fallu simplifier certains aspects de la réalité. Néanmoins, le modèle utilisé est d'un grand réalisme, si on le compare à diverses études de simulation réalisées dans

ce domaine de recherche. Ainsi, ce modèle offre une bonne approximation d'un système réel de production.

CHAPITRE IV

RÉSULTATS: ANALYSE ET DISCUSSION

Le temps de passage moyen des commandes usine a été la mesure de la performance employée lors des différentes simulations comprises dans le design expérimental, et ce en conformité avec l'équation (2,39). Les résultats sont présentés séparément pour une gestion en flux poussés et pour une gestion en flux tirés.

Le niveau quotidien moyen des en-cours associé aux chemins critiques n'a pas été utilisé lors de cette étape. Car le choix des facteurs expérimentaux n'a pas été effectué en fonction de cette mesure de performance. Cette mesure est utilisée essentiellement dans cette recherche pour comparer la performance du système étudié sous une gestion en flux poussés et tirés.

Ensuite, on présente une analyse de l'étude comparative en employant comme mesure de la performance le temps de passage moyen des commandes usine et le niveau quotidien moyen des en-cours associé aux chemins critiques. Ces mesures proviennent des essais de validation effectués à l'aide des modèles obtenus

précédemment en flux poussés et tirés. Notons que ces essais de validation sont nécessaires car les méthodes Taguchi ont été utilisées lors de l'élaboration du design expérimental.

Résultats en flux poussés

Les temps de passage moyens des commandes usine obtenus lors des essais sont présentés au Tableau 9. La moyenne et l'écart-type de chaque facteur expérimental sont présentés au Tableau 10, et le Tableau 11 montre l'analyse de la variance.

Tableau 9

Résultats de simulation en flux poussés: temps de passage
moyen des commandes usine

Numéro d'essai	Temps de passage moyen (jour)		
1	7,08	7,11	6,82
2	7,08	7,40	6,91
3	6,90	7,47	7,03
4	6,83	7,31	7,08
5	6,51	6,79	6,94
6	7,49	6,89	6,79
7	7,43	7,43	7,19
8	6,90	6,68	6,59
9	6,40	6,74	6,73
10	5,99	6,37	6,92
11	6,36	6,70	6,09
12	6,12	6,53	6,46

Tableau 9 (suite)

Numéro d'essai	Temps de passage moyen (jour)		
	13	5,57	5,47
14	5,79	5,75	6,15
15	5,42	5,82	5,10
16	5,72	5,90	6,21

Tableau 10

Résultats de simulation en flux poussés: temps de passage moyen des commandes usine (moyenne et écart-type)

Facteur	Niveau un		Niveau deux	
	Moyenne	Écart-type	Moyenne	Écart-type
T_k	7,03	0,28	6,10	0,47
C	6,77	0,41	6,36	0,70
O	6,55	0,68	6,58	0,53
$tpas_{CU}$	6,58	0,56	6,55	0,65
NS	6,51	0,61	6,62	0,60
T_k^C	6,42	0,73	6,71	0,41
T_k^O	6,61	0,56	6,52	0,65
CO	6,60	0,51	6,53	0,70

Tableau 11

Résultats de simulation en flux poussés: temps de passage
moyen des commandes usine (analyse de la variance)

Source	DF	SS	F	Pr > F
Modèle	8	13,581	18,54	0,0001
Erreur	39	3,571		
Total	47	17,152		
$R^2 = 0,79$				
T_k	1	10,277	112,23	0,0001
C	1	1,989	21,72	0,0001
O	1	0,008	0,09	0,7654
tpascu	1	0,006	0,07	0,7944
NS	1	0,142	1,55	0,2206
$T_k C$	1	1,018	11,12	0,0019
$T_k O$	1	0,095	1,03	0,3159
CO	1	0,046	0,51	0,4815

On remarque au Tableau 11 que la durée cible d'un ensemble d'opérations T_k utilisée aux équations (2,12) et (2,14) affecte significativement le temps de passage moyen des commandes usine. Et il en va de même pour le coefficient appliqué sur les temps de mise en course aux goulots 0 et 1. De plus, l'interaction de ces deux facteurs affecte également le système. Néanmoins, les autres facteurs employés n'affectent pas significativement celui-ci. Ce qui est une bonne nouvelle.

Ainsi, le système de production n'est pas très sensible au choix des règles d'ordonnancement utilisées. On peut donc utiliser l'un ou l'autre des choix proposés. De plus, l'utilisation d'un niveau de service recherché n'affecte pas le temps de passage moyen des commandes usine. Conséquemment, on peut diminuer les pénuries de stock sans affecter le temps de passage. Enfin, la variation de la borne tp_{ascu} de l'équation (2,36) n'a pas affecté la performance du système à l'étude.

Le Tableau 12 montre la régression du modèle ajusté provenant de l'analyse de la variance. Notons que le paramètre T_k a une valeur égale à l'infini au niveau un. Afin de simplifier le calcul de la régression, on a remplacé cette valeur par un nombre très grand, soit 100 heures. Puisque ce modèle est représentatif, il a été utilisé lors des essais de validation en flux poussés.

Tableau 12

Temps de passage moyen des commandes usine en flux
poussés: régression (modèle ajusté)

Source	DF	Moyenne au carré	F	Pr > F
Régression	3	4,426	50,317	0,0001
Erreur	44	0,088		
$R^2 = 0,77$				
Paramètre estimé	DF	Coefficient	t	Pr > t
Ordonnée à l'origine	1	4,986	25,125	0,0001
T_k	1	0,019	6,655	0,0001
C	1	1,444	5,753	0,0001
$T_k C$	1	-0,012	-3,417	0,0014

Résultats en flux tirés

Le Tableau 13 présente les temps de passage moyens des commandes usine obtenus, et on retrouve au Tableau 14 la moyenne ainsi que l'écart-type de chaque facteur. Enfin, l'analyse de la variance est présentée au Tableau 15.

Tableau 13

Résultats de simulation en flux tirés: temps de passage
moyen des commandes usine

Numéro d'essai	Temps de passage moyen (jour)		
1	7,07	7,82	7,06
2	6,99	6,87	6,98
3	6,99	7,35	6,88
4	6,57	6,94	6,70
5	7,25	7,14	7,60
6	7,54	6,87	6,95
7	7,44	7,36	8,03
8	6,69	6,51	6,76
9	5,66	6,35	5,57
10	6,27	6,88	6,74
11	6,31	6,80	7,27
12	5,59	5,67	5,86
13	5,86	5,39	6,08
14	5,76	5,72	5,98
15	5,66	5,06	5,22
16	6,26	5,82	5,81

Tableau 14

Résultats de simulation en flux tirés: temps de passage
moyen des commandes usine (moyenne et écart-type)

Facteur	Niveau un		Niveau deux	
	Moyenne	Écart-type	Moyenne	Écart-type
T _k	7,10	0,39	5,98	0,54
C	6,63	0,60	6,45	0,84

Tableau 14 (suite)

Facteur	Niveau un		Niveau deux	
	Moyenne	Écart-type	Moyenne	Écart-type
O	6,63	0,88	6,45	0,55
tpascu	6,54	0,79	6,54	0,69
NS	6,42	0,69	6,66	0,76
α	6,60	0,71	6,48	0,77
β	6,56	0,80	6,52	0,68
γ	6,45	0,76	6,63	0,70
TC	6,37	0,74	6,71	0,69
TO	6,68	0,76	6,40	0,68
CO	6,58	0,65	6,51	0,82

Tableau 15

Résultats de simulation en flux tirés: temps de passage
moyen des commandes usine (analyse de la variance)

Source	DF	SS	F	Pr > F
Modèle	11	19,508	11,46	0,0001
Erreur	36	5,569		
Total	47	25,077		
$R^2 = 0,78$				
T_k	1	14,930	96,50	0,0001
C	1	0,409	2,64	0,1127
O	1	0,420	2,71	0,1081
tpascu	1	0,001	0,00	0,9448
NS	1	0,723	4,67	0,0374
α	1	0,169	1,09	0,3026
β	1	0,020	0,13	0,7185

Tableau 15 (suite)

Source	DF	SS	F	Pr > F
γ	1	0,405	2,62	0,1143
$T_k C$	1	1,425	9,21	0,0045
$T_k O$	1	0,949	6,14	0,0181
CO	1	0,057	0,37	0,5487

On constate encore l'influence du facteur T_k sur le système en flux tirés, mais pas le coefficient sur les temps de mise en course. Cependant, l'impact de l'interaction de ces deux facteurs demeure.

Les autres facteurs n'affectent pas le système en flux tirés, à l'exception du niveau de service recherché qui varie proportionnellement avec le temps de passage moyen des commandes usine. Ainsi, une augmentation de la durée des temps opératoires a un impact plus important en flux tirés qu'une augmentation des temps de mise en course pour ce système.

On présente au Tableau 16 le modèle ajusté qui a été utilisé lors des essais de validation en flux tirés. Lors de la régression, on a également utilisé un nombre très grand (100 heures) pour le paramètre T_k au niveau un.

Tableau 16

Temps de passage moyen des commandes usine en flux
tirés: régression (modèle ajusté)

Source	DF	Moyenne au carré	F	Pr > F
Régression	4	5,258	24,862	0,0001
Erreur	43	0,211		
$R^2 = 0,72$				
Paramètre estimé	DF	Coefficient	t	Pr > t
Ordonné à l'origine	1	5,547	23,606	0,0001
T_k	1	0,014	4,367	0,0001
NS	1	0,005	2,084	0,0487
T_k^C	1	-0,003	-2,943	0,0052

La validation des modèles et l'analyse comparative

Afin de comparer la performance du système de production sous une gestion en flux poussés et tirés, et afin de valider les modèles obtenus précédemment, un essai a été effectué pour chacun des deux modes de gestion sous des conditions optimales. De plus, chaque essai a été répété cinq fois. Les semences principales utilisées pour la génération des nombres pseudo-aléatoires ont également été choisies à l'aide d'une table de nombres aléatoires.

Afin d'affecter aux différents facteurs les niveaux à employer pour obtenir des conditions optimales, on a observé préalablement les modèles obtenus. Ainsi, on peut déterminer les niveaux pour les facteurs T_k , C et NS à l'aide des Tableaux 12 et 16.

Afin de déterminer les niveaux pour les autres facteurs, un test de Duncan a été effectué. Ces résultats sont présentés aux Tableaux 17 et 18, pour une gestion en flux poussés et tirés respectivement.

Tableau 17

Temps de passage moyen des commandes usine en flux
poussés: test de Duncan

	α	DF	EMS	
	0,05	32	0,0771	
	Groupement		Moyenne	Essai
	A		7,3500	7
B	A		7,1333	3
B	A		7,1300	2
B	A		7,0733	4
B	A		7,0567	6
B	A		7,0033	1
B	C		6,7467	5
B	C		6,7233	8
B	C		6,6233	9
D	C		6,4267	10
D	C	E	6,3833	11
D	C	E	6,3700	12
D	F	E	5,9433	16
	F	E	5,8967	14
	F		5,7233	13
	F		5,4467	15
Essais significativement différents				
7 > 5,8,9,10,11,12,13,14,15,16				
1,2,3,4,6 > 10,11,12,13,14,15,16				
5,8,9,10 > 13,14,15				
11,12,14,16 > 13,15				

Tableau 18

Temps de passage moyen des commandes usine en
flux tirés: test de Duncan

	α	DF	EMS	
	0,05	32	0,0920	
	Groupement		Moyenne	Essai
	A		7,6100	7
B	A		7,3300	5
B	A		7,3167	1
B	A	C	7,1200	6
B	A	C	7,0733	3
B		C	6,9467	2
B		C	6,7933	11
		C	6,7367	4
		C	6,6533	8
		C	6,6300	10
	D		5,9633	16
E	D		5,8600	9
E	D		5,8200	14
E	D		5,7767	13
E	D		5,7067	12
E			5,3133	15
Essais significativement différents				
7 > 2,4,8,9,10,11,12,13,14,15,16				
1,5 > 4,8,9,10,12,13,14,15,16				
3,6 > 9,12,13,14,15,16				
2,4,8,10,11 > 9,12,13,14,15,16				
16 > 15				

Ainsi, on peut déterminer les niveaux pour les facteurs O , $tpas_{CU}$ et γ . Finalement, les niveaux et l'interaction des facteurs α et β ont été choisis aléatoirement. Le Tableau 19 présente une compilation de cet exercice.

TABLEAU 19

Affectation des niveaux utilisés lors des essais
de validation en flux poussés et tirés

Facteur	Niveau	Valeur
T_k	2	3,5
C	2	0,5
O	1	1
$tpas_{CU}$	2	7
NS	2	95
α^\dagger	2	30
β^\dagger	1	25
γ^\dagger	1	20

† Facteurs utilisés en flux tirés uniquement.

Les résultats des simulations pour fins de validation et d'analyse sont présentés aux Tableaux 20 et 21. Les Figures 5. et 6. montrent les résultats obtenus pour le temps de passage moyen des commandes usine versus les modèles ajustés à un intervalle de confiance de 95%. On remarque que seule la quatrième simulation en flux poussés obtient une performance en dehors des limites établies pour cet intervalle de confiance.

Tableau 20

Validation en flux poussés: compilation des résultats

Mesure de la performance	Simulation				
	v-1	v-2	v-3	v-4	v-5
Temps de passage moyen des commandes usine (jour)	5,85	6,36	5,67	7,86	5,72
Nombre quotidien d'en-cours moyen associé aux chemins critiques (un)	1650	1850	1380	1680	1330

Tableau 21

Validation en flux tirés: compilation des résultats

Mesure de la performance	Simulation				
	v-1	v-2	v-3	v-4	v-5
Temps de passage moyen des commandes usine (jour)	5,65	5,77	5,94	5,98	5,65
Nombre quotidien d'en-cours moyen associé aux chemins critiques (un)	724	971	956	823	828

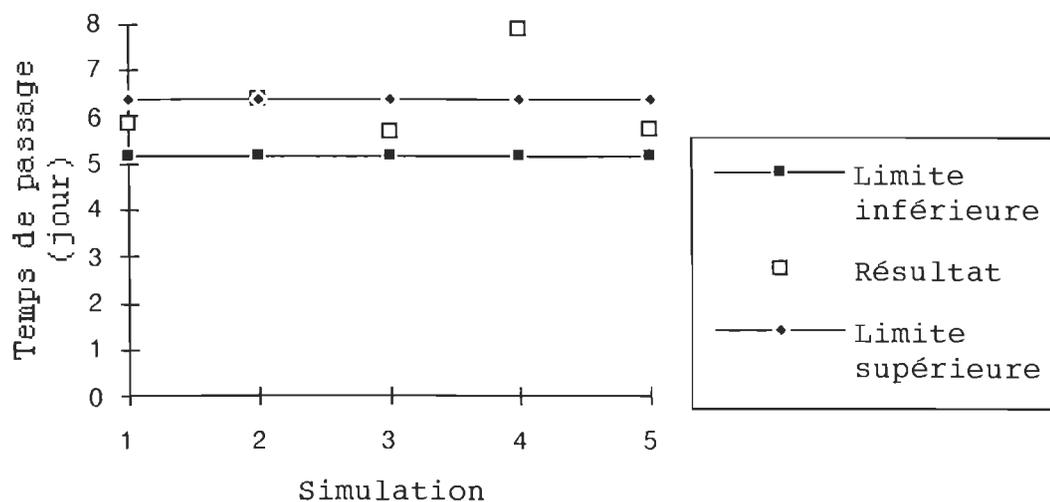


Figure 5. Temps de passage moyen des commandes usine en flux poussés: simulations de validation (intervalle de confiance à 95%)

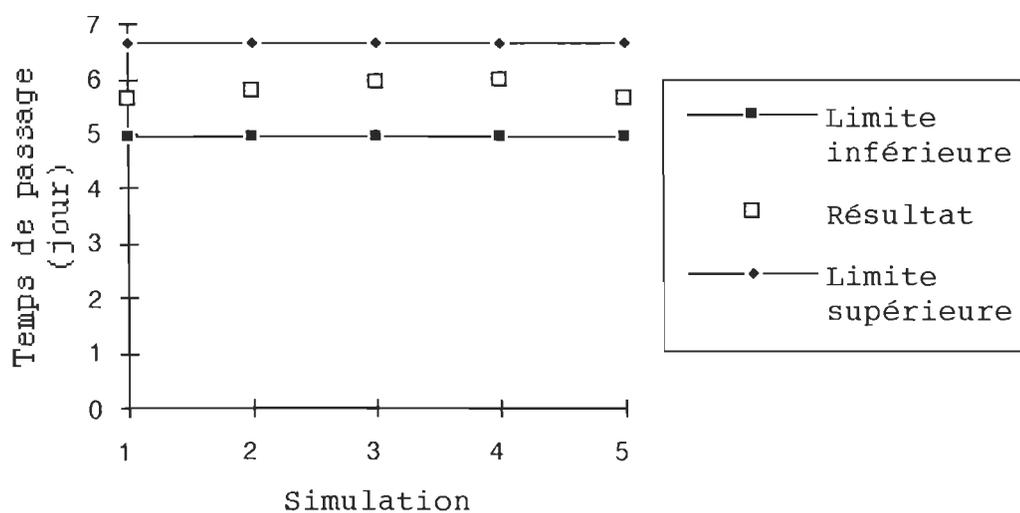


Figure 6. Temps de passage moyen des commandes usine en flux tirés: simulations de validation (intervalle de confiance à 95%)

Ceci est dû au fait qu'au cours de cette simulation, une congestion s'est produite au poste de travail 27 qui est un goulot d'étranglement. Les temps opératoires qui sont générés de façon aléatoires étaient assez longs pendant une certaine période. Comme la planification des ressources est effectuée d'une manière agrégée à l'aide de temps moyens, cette ressource était engorgée. Et puisque le simulateur ne tient pas compte du temps supplémentaire, la file d'attente de ce poste n'a fait que croître. Comme cette situation exceptionnelle ne s'est produite qu'une seule fois, on peut soutenir que la validation est adéquate.

L'analyse comparative poussés versus tirés a été effectuée à l'aide d'un test de Student à variances inconnues et inégales (problème de Behrens-Fisher), à un intervalle de confiance de 95% pour le temps de passage moyen des commandes usine. On observe qu'il n'y a pas de différence significative.

Cependant, il existe une différence significative pour le niveau quotidien d'en-cours moyen associé aux chemins critiques. Dans ce cas, on a effectué un test de Student à variances inconnues mais égales, à un intervalle de confiance de 95%. En fait, le niveau d'en-cours sur les chemins critiques diminue de près de 50% en flux tirés pour ce système de

production. Ce qui est un résultat fort intéressant, surtout compte tenu de l'amplitude de celui-ci.

Notons que pour ces deux tests, on a émis comme hypothèse que les quatre populations étaient normales. Enfin, le Tableau 22 présente un sommaire des résultats de l'étude comparative.

Tableau 22

Résultats de l'étude comparative

Mesure de la performance	Poussés		Tirés		Différence significative
	m	s	m	s	
Temps de passage moyen des commandes usine (jour)	6,29	0,92	5,80	0,16	non
Nombre quotidien d'en-cours moyen associé aux chemins critiques (un)	1578	218	860	103	oui

CHAPITRE V

CONCLUSIONS ET RECOMMANDATIONS

Le but de cette recherche était d'étudier le comportement d'un système réel de production, composé d'un atelier spécialisé qui alimente des lignes d'assemblage, dans un environnement JAT.

Pour y arriver, il a fallu adapter le JAT à ce système en s'inspirant de la méthode CONWIP. Méthode qui utilise essentiellement les goulots d'étranglement pour établir d'une part la taille des commandes usine à lancer en production, et pour déterminer d'autre part les flux qui devront être tirés, soit ceux en amont des goulots.

Le modèle de simulation décrit dans ce mémoire a prouvé son utilité pour atteindre le but recherché. Il a été possible de simuler le système sous une gestion en flux poussés et tirés, et d'établir une base de comparaison valable entre ces deux modes de gestion.

On peut conclure que le JAT peut être applicable à ce type de système de production. Cependant, il ne faut pas s'attendre

nécessairement à une amélioration du temps de passage des commandes usine. Le gain est surtout relié à la diminution du volume d'en-cours sur le plancher de l'usine. Et cette diminution peut être très significative. Ces résultats sont comparables à ceux de Durmusoglu (1991), ainsi qu'à ceux de Ramudhin et Rochette (1991).

Les résultats de cette recherche sont applicables à une multitude d'entreprises manufacturières. Mais il pourrait être difficile pour celles-ci d'établir le nombre de kanbans génériques nécessaires au bon fonctionnement de leur système. Un moyen intéressant pour y parvenir serait l'emploi d'un simulateur accessible pour les petites et moyennes entreprises.

Le progiciel de simulation développé pour cette recherche ne pourrait pas remplir ce rôle adéquatement. En effet, le logiciel SLAM II utilisé pour représenter la partie "usine" de l'entreprise est trop limité côté mémoire. Ainsi, on a été restreint dans le choix des scénarios à simuler à cause de ce logiciel, et non à cause de la capacité des ordinateurs accessibles actuellement sur le marché.

Une première extension de cette recherche serait de bonifier le modèle employé en ajoutant par exemple la possibilité de simuler du temps supplémentaire. Mais il

faudrait reprogrammer le progiciel en utilisant un logiciel de simulation qui offre des perspectives plus intéressantes.

Une seconde extension de cette recherche serait d'analyser la robustesse de la méthode CONWIP, en simulant une multitude de scénarios. Ce qui ne pourrait être fait également qu'à l'aide d'un nouveau progiciel de simulation.

RÉFÉRENCES

1. Ouvrages de références

GIARD, V. (1988) Gestion de la production, 2^e édition, Collection Gestion, Economica.

MONDEN, Y. (1991) Toyota Production System, 2^e édition, Institute of Industrial Engineers.

PRITSKER, A. A.B. (1986) Introduction to Simulation and SLAM II, 3^e édition, John Wiley & Sons.

WANTUCK, K.A. (1989) JUST-IN-TIME For America, KWA Media.

2. Articles de périodiques

BAKER, K.R., POWELL, S.G. et PYKE, D.F. (1990) The performance of push and pull systems: a corrected analysis, International Journal of Production Research, vol. 28, n^o 9: 1731-1736.

CHU, C.-H. et SHIH, W.L. (1992) Simulation studies in JIT production, International Journal of Production Research, vol. 30, n^o 11: 2573-2586.

Elvers, D.A. (1974) The sensitivity of the Relative Effectiveness of Job Shop Dispatching Rules with Respect to Various Arrival Distributions, AIIE Transactions, vol. 6, n^o 1: 41-49.

GOLDRATT, M.E. (1988) Computerized shop floor scheduling, International Journal of Production Research, vol. 26, n^o 3: 443-445.

GOLHAR, D.Y. et STAMM, C.L. (1991) The just-in-time philosophy: A literature review, International Journal of Production Research, vol. 29, n^o 4: 657-676.

KRAJEWSKI, L.J. et al. (1987) Kanban, MRP, and shaping the manufacturing environment, Management Science, vol. 33, n^o 1: 39-57.

PANWALKAR, S.S. et ISKANDER, W. (1977) A Survey of Scheduling Rules, Operations Research, vol. 25, n^o1: 45-61.

RAMASESH, R. (1990) Dynamic Job Shop Scheduling: A Survey of Simulation Research, *International Journal of Management Science*, vol. 18, n° 1: 43-57.

SARKER, B.R. et FITZSIMMONS, J.A. (1989) The performance of push and pull systems: a simulation and comparative study, *International Journal of Production Research*, vol. 27, n° 10: 1715-1731.

SPEARMAN, M.L., WOODRUFF, D.L. et HOPP, W.J. (1990) CONWIP: a pull alternative to kanban, *International Journal of Production Research*, vol. 28, n° 5: 879-894.

SPEARMAN, M.L. et ZAZANIS, M.A. (1992) Push and Pull Production Systems: Issues and Comparisons, *Operations Research*, vol. 40, n° 3: 521-532.

SUGIMORI, Y., et al. (1977) Toyota production system and Kanban system: Materialization of a just-in-time and respect-for-human system, *International Journal of Production Research*, vol. 15: 553-564.

ZÄPFEL, G. et MISSBAUER, H. (1993) New concepts for production planning and control, *European Journal of Operational Research*, vol. 67: 297-320.

3. Comptes rendus

Durmusoglu, M.B. (1991) Comparison of Push and Pull Systems in a Cellular Manufacturing Environment, Proceedings of the International Conference on Just-in-Time Manufacturing Systems-Operational Planning and Control Issues, Montréal, Elsevier Science Publishers B.V.: 115-132.

Ramudhin, A. et Rochette, R. (1991) Just-in-Time Practices in an Unstable Environment: A Simulation Study, Proceedings of the International Conference on Just-in-Time Manufacturing Systems-Operational Planning and Control Issues, Montréal, Elsevier Science Publishers B.V.: 133-140.

4. Rapport interne

POIRIER, P. (1995) Données reliées à la modélisation du système productif de l'entreprise Meubles Laurier Ltée, Compilation finale effectuée dans le cadre du projet C.R.S.N.G.: Simulateur générique, Université du Québec à Trois-Rivières.

ANNEXE A

Modélisation graphique du réseau en SLAM II

Cette section présente brièvement la modélisation graphique du réseau en langage SLAM II employée pour représenter la partie "usine" du simulateur.

La Figure 7. montre les blocs ressources représentant les postes de travail de l'usine. Toutes les ressources ont une capacité d'une unité lors de l'initialisation. Notons que les numéros de fichier sont de priorité décroissante de gauche à droite pour une ressource donnée. Par exemple, si une entité peut être emmagasinée simultanément dans le fichier 101 ou 100 de la ressource MACHIN00, celle-ci sera emmagasinée dans le fichier 101 prioritairement.

La Figure 8. présente l'aiguilleur du réseau. Une entité qui entre dans le réseau par cette branche sera aiguillée vers: (a) DEG, pour décomposer une commande client produits en commandes client produit; (b) EXP, pour les ordres d'expédition afin de reconstituer les commandes clients produits avant l'expédition; (c) STO, pour acheminer les commandes usine complétées vers les stocks de produits finis; (d) TRA, pour

MACHIN00	1	101	100	51	1
MACHIN01	1	102	100	52	2
MACHIN02	1	103	100	53	3
MACHIN03	1	104	100	54	4
MACHIN04	1	105	100	55	5
MACHIN05	1	106	100	56	6
MACHIN06	1	107	100	57	7
MACHIN07	1	108	100	58	8
MACHIN08	1	109	100	59	9
MACHIN09	1	110	100	60	10
MACHIN10	1	111	100	61	11
MACHIN11	1	112	100	62	12
MACHIN12	1	113	100	63	13
MACHIN13	1	114	100	64	14
MACHIN14	1	115	100	65	15
MACHIN15	1	116	100	66	16
MACHIN16	1	117	100	67	17
MACHIN17	1	118	100	68	18
MACHIN18	1	119	100	69	19
MACHIN19	1	120	100	70	20
MACHIN20	1	121	100	71	21

MACHIN21	1	122	100	72	22
MACHIN22	1	123	100	73	23
MACHIN23	1	124	100	74	24
MACHIN24	1	125	100	75	25
MACHIN25	1	126	100	76	26
MACHIN26	1	127	100	77	27
MACHIN27	1	128	100	78	28
MACHIN28	1	129	100	79	29
MACHIN29	1	130	100	80	30
MACHIN30	1	131	100	81	31
MACHIN31	1	132	100	82	32
MACHIN32	1	133	100	83	33
MACHIN33	1	134	100	84	34
MACHIN34	1	135	100	85	35
MACHIN35	1	136	100	86	36
MACHIN36	1	137	100	87	37
ASSEMB37	1	138	88	38	
ASSEMB38	1	139	89	39	
ASSEMB39	1	140	90	40	
ASSEMB40	1	141	91	41	

Figure 7. Les blocs ressources

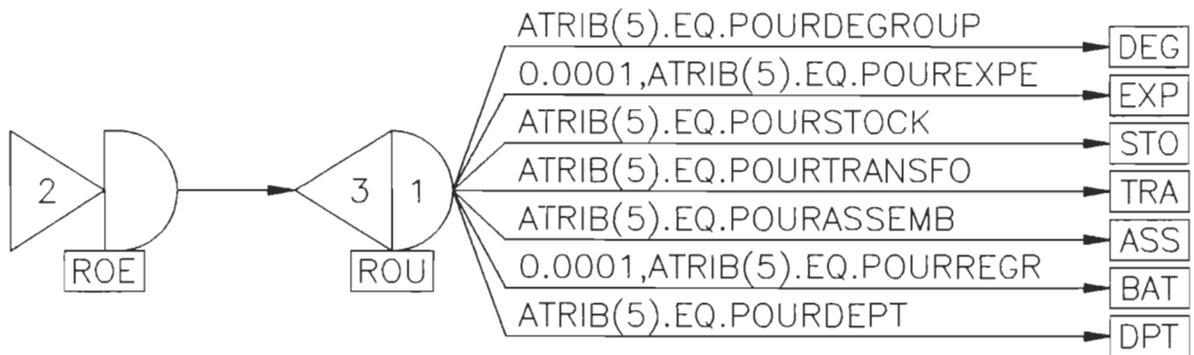


Figure 8. L'aiguilleur du réseau

acheminer un ordre de fabrication à une ressource de transformation; (e) ASS, pour acheminer un ordre d'assemblage à une ressource d'assemblage; (f) BAT, pour regrouper tous les composants nécessaires à la réalisation d'un ordre d'assemblage; et (g) DPT, pour acheminer un lot qui sort d'un centre de travail vers son nouveau centre.

On retrouve à la Figure 9. les branches reliées à l'avancement du temps en cours de simulation. La première branche génère une entité qui appelle le noeud EVENT,1,1; au début d'un horizon de planification afin de générer toute la demande pour le prochain horizon de planification.

La deuxième branche génère une entité qui appelle le noeud EVENT,2,;. et le noeud EVENT,17,;. Le premier noeud sert à appeler au début d'une journée les commandes clients déjà générées afin de simuler une arrivée journalière des commandes. Le second noeud sert à appeler les prises en stock quotidiennes. Une prise en stock sert à piger dans les stocks de produits finis afin de compléter les commandes clients. Elle a lieu uniquement lorsque le niveau physique du stock visé est suffisant pour répondre à la commande client produit la plus prioritaire. L'indice de priorité est établi à l'aide de la règle du premier-arrivé, premier-servi.

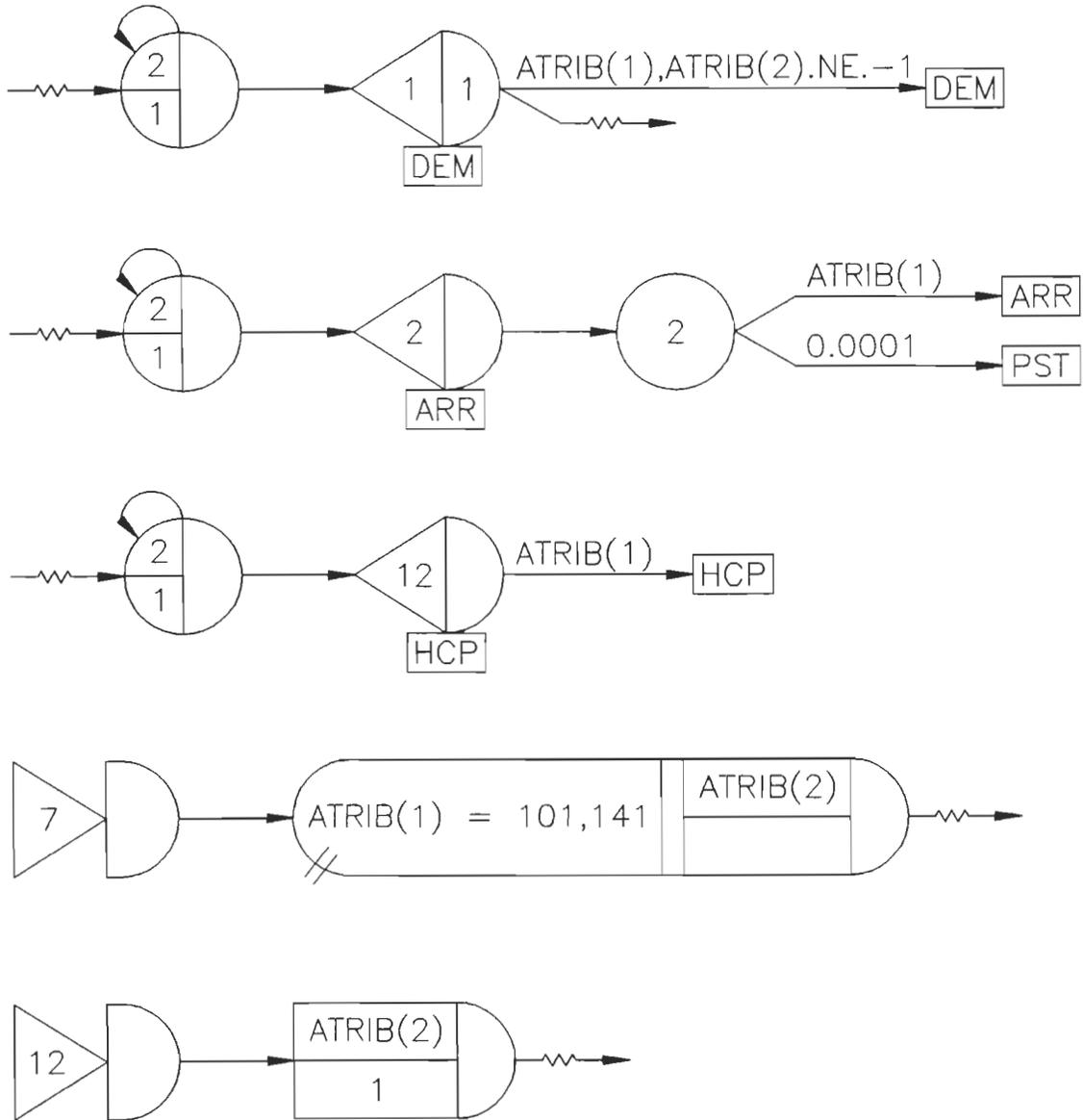


Figure 9. Les branches reliées à l'avancement du temps

Les autres branches servent à arrêter la production aux différents postes de travail à la fin d'une journée selon la capacité horaire journalière de chaque centre de charge. Une entité appelle aux 30 minutes le noeud EVENT,12,;. Cette procédure vérifie la capacité des centres de charge. Lorsqu'une fin de journée est constatée pour un centre de charge donné, un droit de préemption est appliqué sur tous les postes de travail de ce centre, afin d'arrêter la production. La préemption est levée au début de la journée subséquente avec la dernière branche.

La Figure 10. montre les branches servant à la gestion des transferts de lots d'un centre de travail à un autre. Lorsqu'un lot doit entrer dans un nouveau centre, il est mis en attente le temps de vérifier si un kanban générique est disponible pour le transfert. Lorsqu'on lui affecte un kanban, on lui retire le kanban déjà affecté afin de le rendre disponible de nouveau. Notons que les kanbans génériques sont utilisés sous une gestion en flux tirés et poussés. Cependant, la quantité de kanbans génériques disponibles sous une gestion en flux poussés est égale à un nombre très grand.

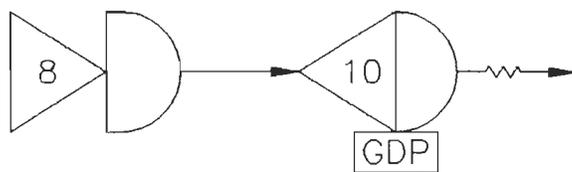
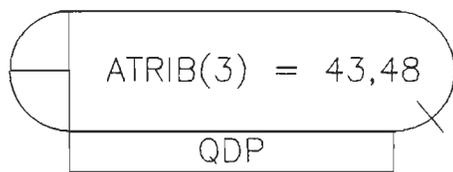
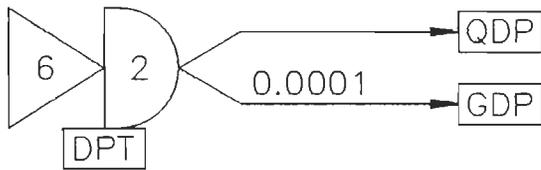


Figure 10. La gestion intercentre de travail

La gestion du niveau des stocks physiques est représentée par la Figure 11. La première branche sert à emmagasiner les demandes de réservation pour chaque produit géré pour stock. La deuxième branche est utilisée pour la prise en stock et la dernière branche sert à détruire l'entité associée à une commande usine complétée après le réajustement du stock visé.

La Figure 12. montre la branche employée pour le regroupement des composants avant l'assemblage. Le noeud `EVENT,15,;` sert à la collecte de statistiques. Les entités sont mises en attente par le noeud `BATCH,5000/3,TRIB(2),,FIRST,NONE,;`. Lorsque tous les composants d'un assemblage sont regroupés, un ordre d'assemblage est lancé à l'aide du noeud `EVENT,8,;`.

On retrouve à la Figure 13. la partie du réseau dédiée à la gestion des postes de travail d'assemblage. Les ordres d'assemblage sont emmagasinés à l'aide du noeud `AWAIT(TRIB(2) = 38,41),TRIB(2),,;`. Lorsqu'une ressource est disponible et qu'au moins une entité est en attente pour celle-ci, l'entité la plus prioritaire appelle la procédure d'ordonnancement via le noeud `EVENT,5,;` et cette entité retourne dans sa file d'attente. Un nouvel ordonnancement est effectué. L'entité la plus prioritaire est enlevée de la première branche pour être insérée dans la deuxième où est simulé l'assemblage.

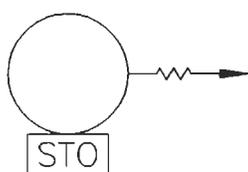
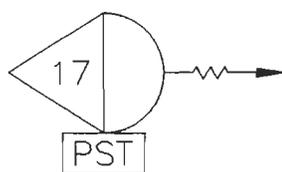


Figure 11. La gestion du niveau physique des stocks

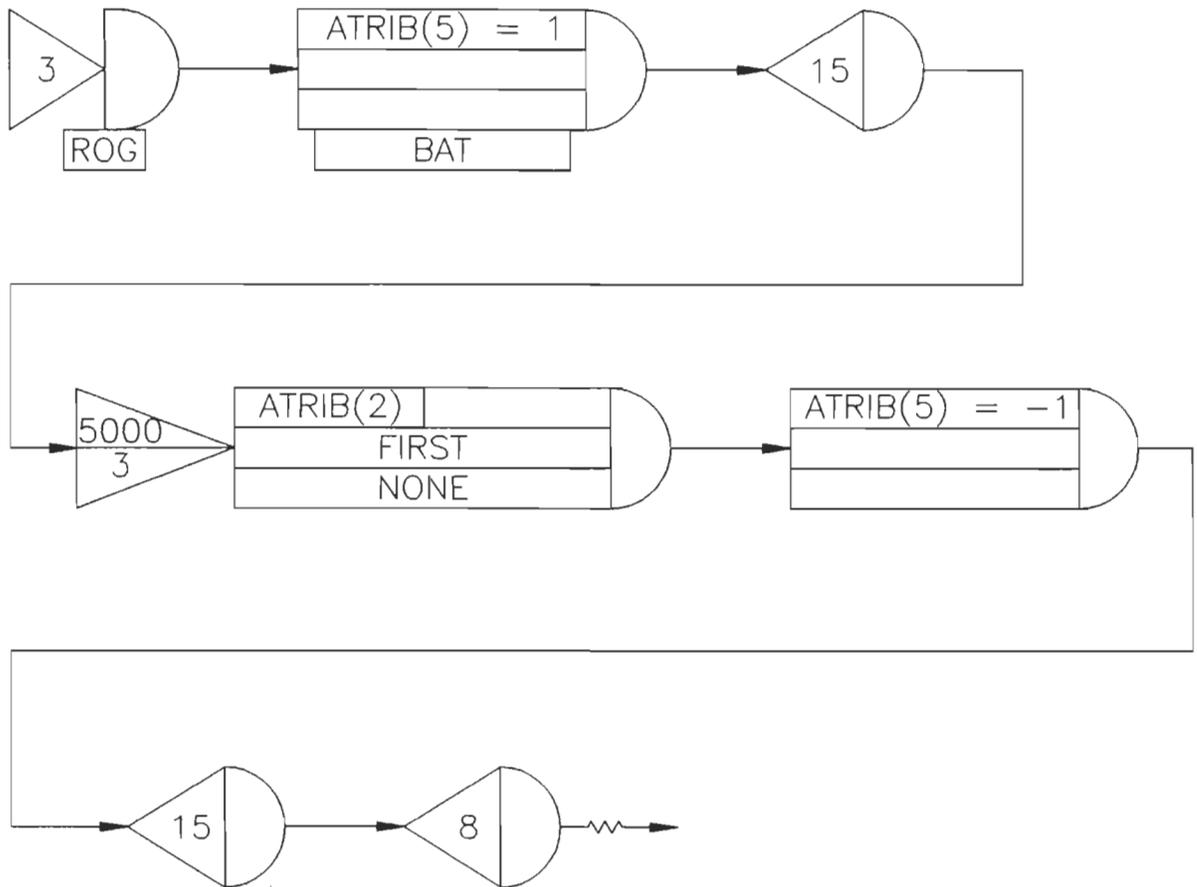


Figure 12. Le regroupement des composants avant l'assemblage

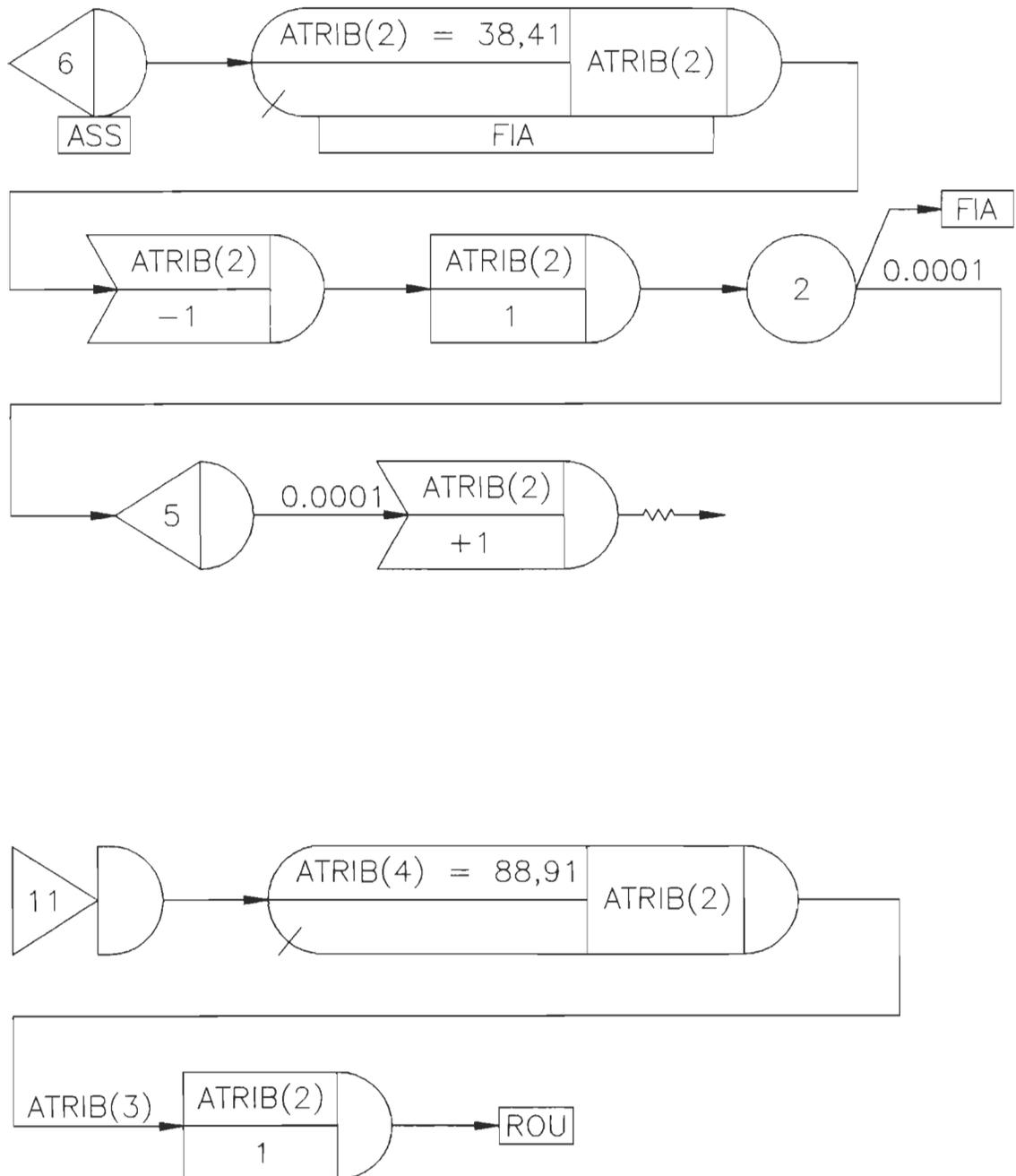


Figure 13. Les ressources d'assemblage

La Figure 14. montre la partie du réseau utilisée pour la représentation des postes de travail de l'atelier spécialisé. La première branche a une fonction similaire à celle de la Figure 13. Cependant, l'entité qui traverse la branche n'appelle pas la procédure d'ordonnancement, elle appelle plutôt la procédure de gestion du personnel.

S'il n'y a pas assez d'ouvriers disponibles, l'entité ayant visité le noeud `EVENT,9,;` est directement détruite. Une nouvelle entité est créée et acheminée vers le noeud `ENTER,14,;` où elle sera dupliquée. La première entité ainsi obtenue est dirigée vers une file d'attente plus prioritaire que les files de la première branche, et la seconde vers le noeud `GOON,;` afin de remettre la capacité de la ressource à un. À ce moment, l'entité dans le fichier le plus prioritaire commence à utiliser la ressource sans ouvrier et avec un temps opératoire indéterminé dans la file 99.

On peut noter ici que les ressources dont leur entité sont dans la file 99 ne peuvent être arrêtées à la fin d'une journée de travail par le noeud `PREEMPT` décrit antérieurement. Néanmoins, ceci n'engendre pas de situations non désirées. Car ces ressources sont considérées en attente d'ouvriers. Elles sont donc inactives.

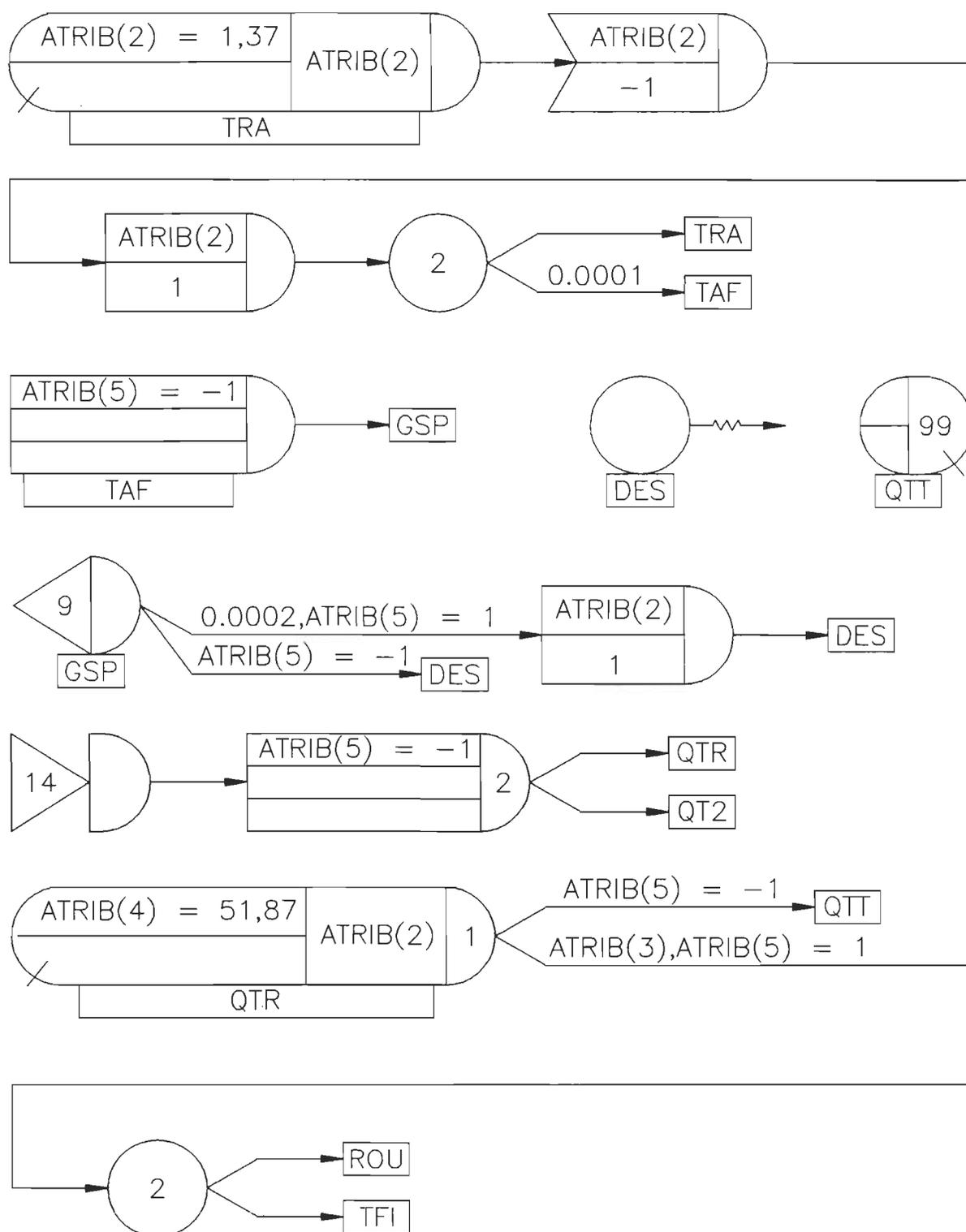


Figure 14. Les postes de travail de l'atelier spécialisé

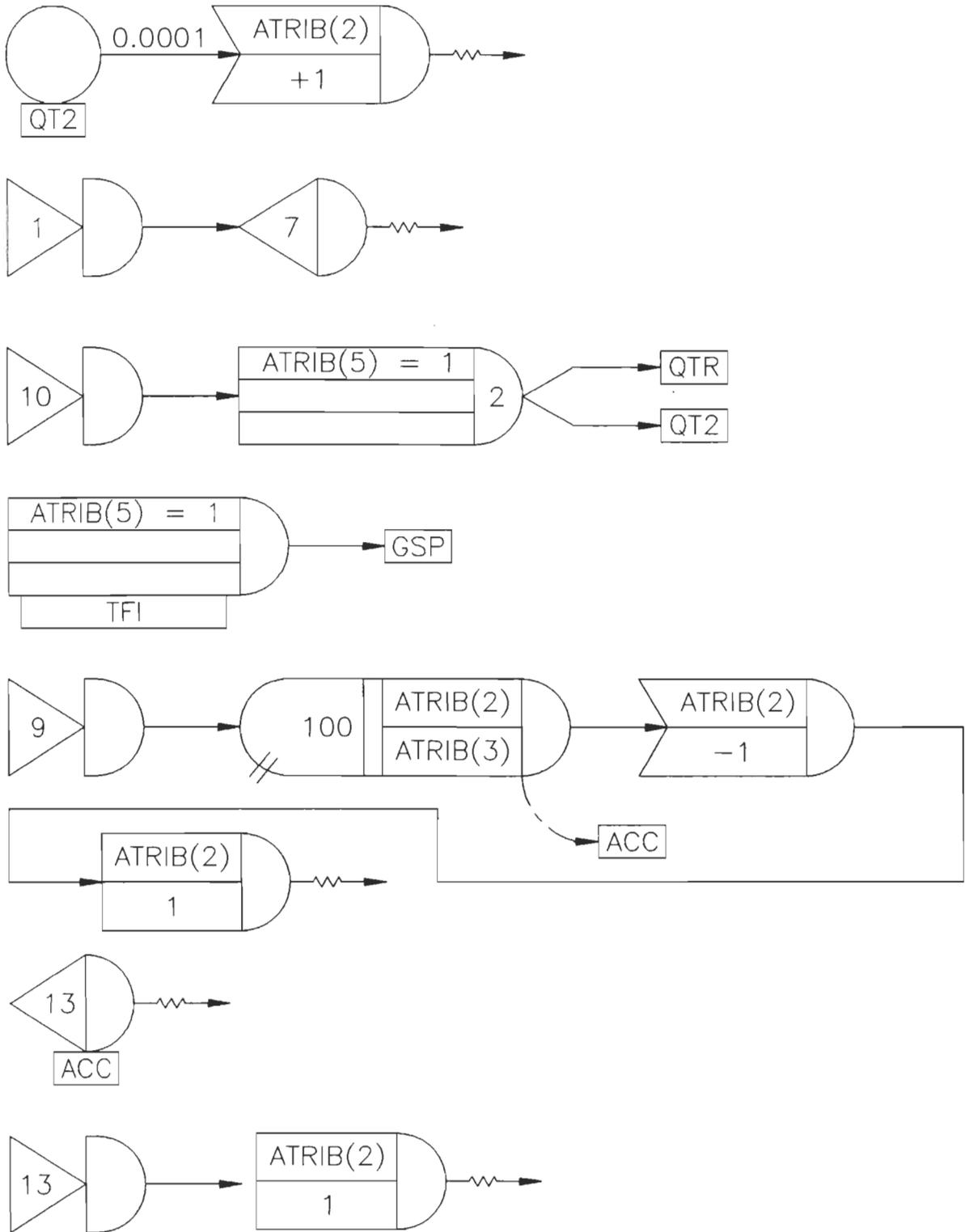


Figure 14. (suite)

S'il y a assez d'ouvriers disponibles, on affecte le minimum d'ouvriers nécessaires s'il y a au moins une entité en attente dans la file 99. Sinon, on affecte à la ressource le plus grand nombre possible d'opérateurs. L'entité ayant visité le noeud EVENT,9,; est directement détruite. Une nouvelle entité est créée et acheminée vers le noeud ENTER,1,;. Elle traverse le noeud EVENT,1,; qui appelle la procédure d'ordonnancement. Cette entité est ensuite détruite. Après l'ordonnancement, une nouvelle entité est générée et acheminée vers le noeud ENTER,10,; où elle sera dupliquée. La première entité ainsi obtenue est dirigée vers une file d'attente plus prioritaire que les files de la première branche, et la seconde vers le noeud GOON,; afin de remettre la capacité de la ressource à un. À ce moment, l'entité dans le fichier le plus prioritaire commence à utiliser la ressource avec un temps opératoire qui est fonction du nombre d'opérateurs affectés.

Lorsqu'une ressource se libère, l'entité est dupliquée. La première entité est acheminée à l'aiguilleur afin de déterminer où se rendra le lot en question. La seconde entité appelle la procédure de gestion du personnel, puis elle libère la ressource. On observe s'il y a au moins une entité en attente dans la file 99. Si oui, un ordonnancement de cette file est effectué et l'entité la plus prioritaire est enlevée de la file puis acheminée vers le noeud ENTER,13,;.

S'il n'y a pas d'entité dans la file 99, on vérifie toute les ressources utilisées à ce moment, par ordre de priorité, afin de détecter si l'on pourrait accélérer le temps opératoire d'une ressource en lui affectant encore plus d'opérateurs. Si c'est le cas, une nouvelle entité est créée et envoyée vers le noeud ENTER,9,;. Le temps opératoire est recalculé (EVENT,13,;) et l'entité est réacheminée vers le noeud AWAIT(ATRIB(4) = 51,87),ATRIB(2),,;. Enfin, une nouvelle entité est créée (EVENT13,;) et dirigée vers le noeud GOON,; afin d'augmenter la capacité de la ressource à un.

La Figure 15. montre la branche utilisée pour décomposer une commande client produits en commandes client produit en appelant la procédure via le noeud EVENT,14,;. Cette figure montre également la branche employée pour la constitution des commandes clients avant l'expédition. Les noeuds EVENT,16,; et EVENT,11,; servent à la collecte de statistiques.

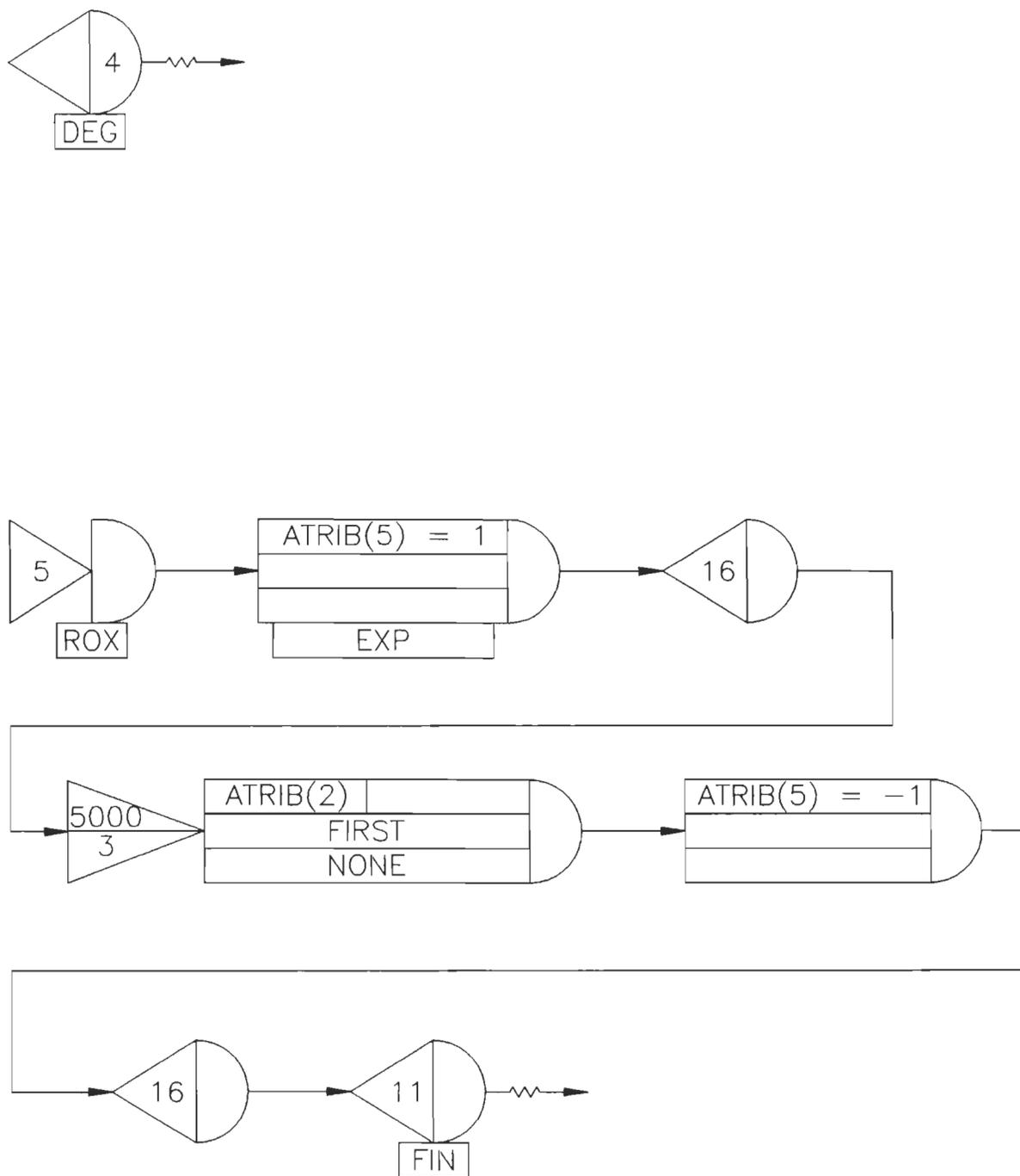


Figure 15. L'éclatement des commandes clients et leur reconstitution

ANNEXE B

Le facteur de sécurité

La Figure 16. présente un exemple d'utilisation du facteur de sécurité de l'équation (2,2). À partir d'une demande réelle, on génère une demande prévisionnelle selon une distribution uniforme. Cette demande prévisionnelle peut varier à l'intérieur de l'enveloppe délimitée par les courbes DP, limite sup. et DP, limite inf. ($PV = 20$).

On utilise le facteur de sécurité pour majorer la demande prévisionnelle afin de réduire les pénuries. La demande prévisionnelle majorée peut varier à l'intérieur de l'enveloppe délimitée par les courbes $FS*DP$, limite sup. et $FS*DP$, limite inf. ($NS = 95$). On utilise cette dernière lors du calcul de la quantité totale à fabriquer et pour établir la position de l'inventaire.

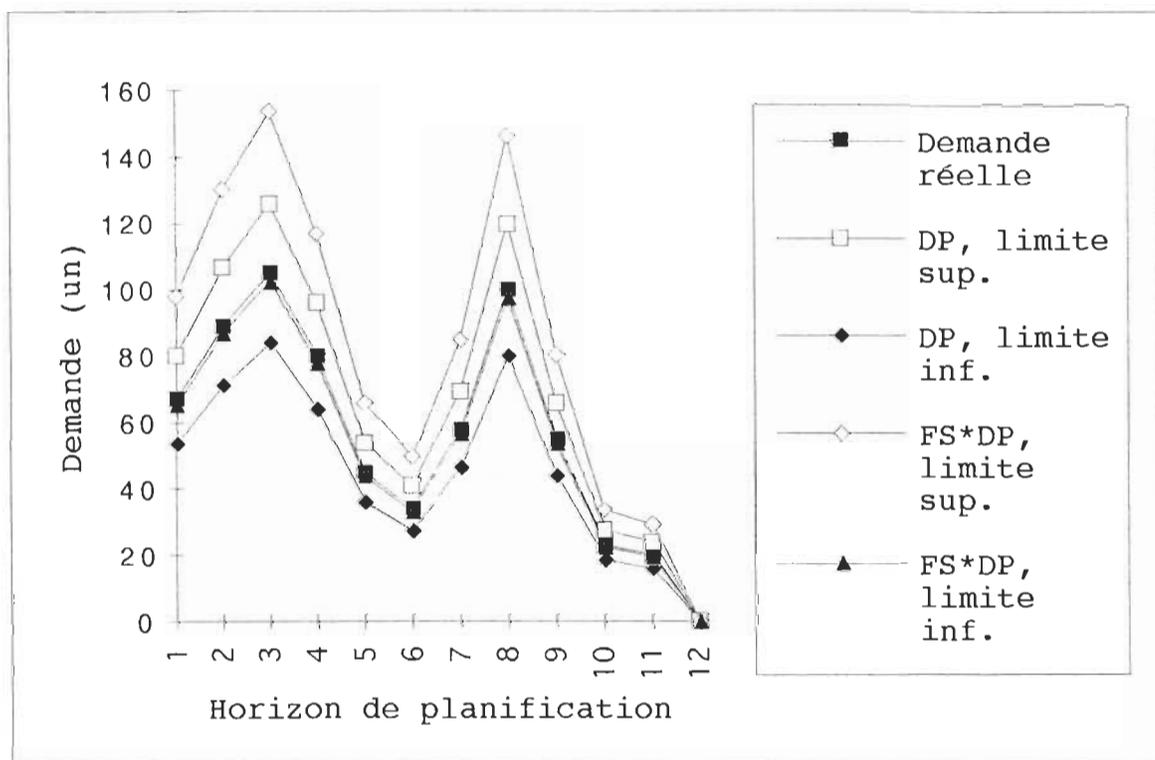


Figure 16. Le facteur de sécurité utilisé pour la gestion des stocks: exemple